

**Evolution komplexer Algen und
der Übergang von Grünalgen zu terrestrischen Pflanzen –
Molekulare und phylogenetische Charakterisierung der Gene des
Primärstoffwechsels in Haptophyten und Streptophyten**

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von René Teich
aus Salzwedel

1. Referent: Professor Dr. Rüdiger Cerff

2. Referent: Professor Dr. Ralf-Rainer Mendel

eingereicht am: 04.09.2006

mündliche Prüfung (Disputation) am: 08.01.2007

Druckjahr 2007

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Petersen, J., **Teich, R.**, Brinkmann, H. & Cerff, R.: A “green” phosphoribulokinase in complex algae with red plastids: evidence for a single secondary endosymbiosis leading to haptophytes, cryptophytes, heterokonts and dinoflagellates. *Journal of Molecular Evolution* 62: 143–157 (2006).

Petersen, J., **Teich, R.**, Becker, B., Cerff, R. & Brinkmann, H.: The GapA/B gene duplication marks the origin of Streptophyta (charophytes and land plants). *Molecular Biology and Evolution* 23: 1109-1118 (2006).

Teich R., Zauner S., Brinkmann H., Petersen J.: Origin and Distribution of Calvin Cycle Fructose and Sedoheptulose Bisphosphatase in Plantae and Complex Algae. A Single Secondary Origin of Complex Red Plastids and Subsequent Propagation via Tertiary Endosymbioses. (Protist in press).

Tagungsbeiträge

Teich, R., Petersen, J. & Cerff, R.: Die Evolution der Haptophyten – Über den Grundstoffwechsel zum Ursprung komplexer Rotalgen. (Vortrag) 9. Jahrestagung der Deutschen Sektion der Internationalen Gesellschaft für Endocytobiologie (ISE-G), Blaubeuren (2003).

Teich, R., Grauvogel, C., Cerff, R. & Petersen, J.: The Phylogenetic Position of Haptophytes According to Genes for the Primary Metabolism. (Poster) Internationale Botanikertagung, Braunschweig (2004).

Teich, R., Petersen, J. & Cerff, R.: GAPDH-Genduplikationen: Evolution der Genstruktur in Charophyceen und Landpflanzen. (Vortrag) 10. Tagung der Sektion Phykologie in der Deutschen Botanischen Gesellschaft (DGB), Königswinter (2005).

Widmung

Diese Arbeit widme ich meiner Großmutter, die bis zuletzt meinen beruflichen und persönlichen Werdegang mit großem Interesse verfolgte und mich stets in allen Bereichen des Lebens vorbehaltlos unterstützte.

Inhaltsverzeichnis

1.	Einleitung	1
1.1	Primäre Endosymbiose und die Entstehung der Landpflanzen	2
1.2	Die Evolution eukaryotischer Algen – sekundäre und tertiäre Endosymbiosen	4
1.3	Algenabteilungen mit komplexen Rhodoplasten	5
1.3.1	Haptophyta	5
1.3.2	Cryptophyta	6
1.3.3	Heterokontophyta (inklusive Diatomeen)	7
1.3.4	Dinophyta	8
1.4	Der Primärmetabolismus photosynthetischer Eukaryoten	10
1.4.1	Photosynthese und Calvin-Zyklus	10
1.4.2	Glykolyse	12
1.5	Ursprung und Evolution der Gene für Glykolyse und Calvin-Zyklus	13
1.6	Spleißosomale Introns – Zeugnisse eukaryotischer Evolution	14
1.7	Ziele der Arbeit	16
 Teil A – Die Evolution der Streptophyta		
2.	Ergebnisse Teil A - Die Evolution der Streptophyta	17
2.1	Etablierung und Analyse plastidärer GAPDH-Sequenzen	18
2.1.1	Etablierung plastidärer GAPDH-Sequenzen mittels degenerierter Primer und der RACE-Technik	18
2.1.2	Phylogenetische Analyse plastidärer GAPDH-Sequenzen	19
2.2	Die Struktur von GAPDH-Genen	23
2.2.1	GapA/B (EC-Nummer 1.2.1.13)	23
2.2.2	GapC/Cp (EC-Nummer 1.2.1.12)	26
 3. Diskussion Teil A - Die Evolution der Streptophyta		
3.1	Die phylogenetische Stellung von <i>Mesostigma</i>	29
3.2	Intron/Exon Struktur von GAPDH-Genen	30
3.2.1	Intronverteilung in GAPDH-Genen der Viridiplantae	30

Inhaltsverzeichnis

3.2.2	GAPDH-Introns: „Neuerwerb“ vor 900 Millionen Jahren oder Relikte der uralten Genstruktur	32
3.2.2.1	Vergleich der Intronpositionen auf Abteilungs-Ebene	33
3.2.2.2	Vergleich der Intronpositionen von <i>GapA</i> und <i>GapC</i> -Genen	33

Teil B – Die Evolution komplexer Algen

4.	Ergebnisse Teil B - Die Evolution komplexer Algen	35
4.1	<i>In silico</i> Arbeiten	35
4.1.1	Taxonsampling für phylogenetische Analysen	35
4.1.2	Authentizität	36
4.1.3	Lokalisierung	36
4.2	Etablierung von Sequenzen	36
4.2.1	EST-Projekt von <i>Pavlova lutheri</i>	36
4.2.2	Degenerierte Primer	37
4.2.3	Homologe Gensonden	38
4.3	Phylogenetische Untersuchungen von Genen aus komplexen Algen mit Rhodoplasten	40
4.3.1	Triosephosphat Isomerase (TPI)	40
4.3.1.1	Phylogenetische Analyse der TPI	41
4.3.2	Enolase (ENO)	43
4.3.2.1	Phylogenetische Analyse der Enolase	43
4.3.3	Fructose-1,6-bisphosphat Aldolase (FBA Klasse I)	47
4.3.3.1	Phylogenetische Analyse der FBA Klasse I	48
4.3.4	Phosphoribulokinase (PRK)	50
4.3.4.1	Charakterisierung der Genstruktur von PRK-Sequenzen	51
4.3.4.2	Phylogenetische Analyse der PRK	53
4.3.5	Fructose-1,6-bisphosphatase (FBP) und Sedoheptulose-1,7-bisphosphatase (SBP)	55
4.3.5.1	Fructose-1,6-bisphosphatase (FBP)	55
4.3.5.1.1	Phylogenetische Analyse der FBP	56
4.3.5.2	Sedoheptulose-1,7-bisphosphatase (SBP)	59
4.3.5.2.1	Genstruktur der SBP	59
4.3.5.2.2	Phylogenetische Analyse SBP	62

Inhaltsverzeichnis

5.	Diskussion Teil B - Die Evolution komplexer Algen	65
5.1	Triosephosphat Isomerase (TPI)	66
5.2	Enolase (ENO)	67
5.3	Fructose-1,6-bisphosphat Aldolase (FBA)	69
5.3.1	FBA Klasse II	69
5.3.2	FBA Klasse I	70
5.3.3	Funktionelle Redundanz von Aldolasen?	71
5.4	Die Phosphoribulokinase (PRK) – ein spezifisches Enzym des Calvin-Zyklus	71
5.5.	Fructose-1,6-bisphosphatase (FBP) und Sedoheptulose-1,7-bisphosphatase (SBP)	74
5.6	Rückschlüsse der Einzelgen-Phylogenien auf die organismische Evolution komplexer Algen	77
5.6.1	Bedeutung der Einzelgenanalysen für die Entstehung komplexer Algen	77
5.6.2	Die komplexe Evolution komplexer Algen - Tertiäre Endosymbiosen	79
6.	Zusammenfassung	81
7.	Material und Methoden	83
7.1	Herkunft des Pflanzenmaterials, der DNA bzw. RNA	83
7.2	<i>Escherichia coli</i> Stämme	85
7.3	Plasmid- und Bakteriophagen Vektoren	85
7.4	Enzyme	86
7.5	Reagenzsätze (Kits)	86
7.6	Sonstiges	87
7.7	Chemikalien	87
7.8	Medien für <i>E. coli</i> und Lambda-Phagen	87
7.9	Geräte	88
7.10	Arbeiten mit <i>E. coli</i>	89
7.11	Arbeiten mit Nukleinsäuren	89
7.11.1	Standardmethoden	89
7.11.2	RNA- und DNA-Isolation	90

Inhaltsverzeichnis

7.11.3	<i>In vivo</i> Exzision rekombinanter λ ZAPII-Phagen	92
7.11.4	PCR-Techniken	93
7.11.4.1	RACE (rapid amplification of cDNA ends)	93
7.11.4.2	Reverse Transkriptase PCR (RT-PCR)	93
7.11.4.3	PCR mit genomischer DNA	93
7.11.5	Sequenzierung	93
7.12	<i>In silico</i> Arbeiten	94
7.12.1	Datenbanken	94
7.12.2	BLAST-Programme	95
7.12.3	Vorhersageprogramme für die intrazelluläre Lokalisierung von Proteinen	95
7.12.4	Phylogenie-Programme	95
7.12.5	Phylogenetische Analysen	95
7.13	Primerliste	96
8.	Literaturverzeichnis	104
9.	Anhang	117
9.1	Accession Numbers	117
9.1.1	Accession Numbers Teil A (cDNA Klone)	117
9.1.2	Accession Numbers Teil A (PCR mit genomischer DNA)	117
9.1.3	Accession Numbers Teil B	117
9.2	Sequenzen	119
9.2.1	<i>Pavlova lutheri</i> Triosephosphat Isomerase (TPI; cDNA Klon)	119
9.2.2	<i>Pavlova lutheri</i> Enolase (cDNA Klon)	121
9.2.3	<i>Pavlova lutheri</i> Fructose-1,6-bisphosphat Aldolase (FBA Klasse I; cDNA Klon)	124
9.2.4	Phosphoribulokinase (PRK) <i>Hanusia phi</i> und <i>Guillardia theta</i>	126
9.2.5	<i>Prymnesium parvum</i> Fructose-1,6-bisphosphatase (FBP; RT-PCR Klon)	128
9.3	Abbildungen	129
	Abbildung 22	129
	Abbildung 23	130
9.4	Abkürzungen	131

1. Einleitung

„Doch außerdem ist im Genom jedes Individuums die Geschichte seiner Art festgehalten, ein ununterbrochener Faden der Abstammung, der bis zu den Anfängen des Lebens zurückreicht. Über diesen Faden, der den Besitz gemeinsamer Gene bedeutet, ist jede Art mit jeder anderen verbunden – je nach Anzahl der geteilten Gene eng oder auch sehr entfernt“

Roger Lewin 1998

Mit der Erkenntnis, dass alle Prozesse des Lebens von vererbten genetischen Informationen gesteuert werden, sahen Biologen auch das Potential für die Rekonstruktion der Evolution. Die Basis für die Erforschung von Verwandtschaftsbeziehungen wechselte so von klassischen phänotypischen Kriterien, wie der Morphologie oder biochemischen Eigenschaften, auf eine molekulare Ebene. Anfänglich verglich man immunologische Merkmale beispielsweise bestimmter Blutproteine des Menschen und einiger Primaten zur Klärung des Verwandtschaftsgrades (Zuckerland et al. 1960). Nach der Etablierung von Sequenzierungsmethoden für Proteine und Gene stand ein „chirurgisches“ Werkzeug zur gezielten Klärung evolutiver Fragen zur Verfügung. 1990 definierten Woese und Mitarbeiter anhand von 16S rDNA Analysen die drei Domänen des Lebens – Bacteria, Archaea und Eukarya. Diese Entdeckung revolutionierte die traditionelle Einteilung der Organismen in die fünf Reiche Monera (Bakterien), Pflanzen, Tiere, Pilze und Protisten (Whittaker 1959) und verdeutlichte, dass die Aufspaltung in Tiere und Pflanzen in der Evolution erst relativ spät erfolgte. Die erfolgreiche Aufklärung solcher Ereignisse ist stark von der Wahl der untersuchten Gene abhängig, die mit einer spezifischen Frequenz Mutationen ansammeln (Konzept der molekularen Uhr). Lange zurückliegende Aufspaltungen lassen sich mit eher langsam evolvierenden (= stark konservierten) Markern untersuchen, wohingegen schnell mutierende DNA-Bereiche jüngere Ereignisse auflösen. Die Darstellung dieser so genannten phylogenetischen Analysen erfolgt meist in Form eines Baumes, da ein solcher die vertikale Evolution, d.h. die Weitergabe des Erbguts an die Nachkommen und damit die stete Aufspaltung der Arten, widerspiegelt. Auf organismischer Ebene prägten darüber hinaus seltene horizontale Ereignisse wie die Verschmelzung unabhängiger Organismen durch Endosymbiosen die Evolution der Eukaryoten entscheidend. Die Aufklärung daraus resultierender vernetzter Verwandtschaftsbeziehungen ist das Ziel der modernen Evolutionsforschung.

1. Einleitung

1.1 Primäre Endosymbiose und die Entstehung der Landpflanzen

Heutige photoautotrophe Eukaryoten lassen sich auf eine einzige Endosymbiose mit einem Cyanobakterium zurückführen (Rodríguez-Ezpeleta et al. 2005), welches von einer heterotrophen Wirtszelle aufgenommen und sukzessive zum Plastiden reduziert wurde. Aus dieser Endosymbiose gingen die drei primären photosynthetischen Linien der Glaucocystophyta, Rotalgen (Rhodophyta) sowie Grünalgen (Chlorophyta) und Landpflanzen hervor (siehe Abb. 1), die jeweils zwei Plastidenmembranen besitzen. Aufgrund ihres gemeinsamen Ursprungs werden Grünalgen und Landpflanzen auch als Viridiplantae („grüne Pflanzen“) zusammengefaßt, während alle drei Hauptlinien zusammen als „Plantae“ bezeichnet werden (Bhattacharya und Medlin 1995; Douglas 1998; Rodríguez-Ezpeleta et al. 2005). Die Glaucocystophyta werden angesichts primitiver morphologischer Merkmale wie den Cyanellen (ursprüngliche Plastiden), die ebenso wie Cyanobakterien von einer Peptidoglycan-Wand umgeben sind, und Carboxysomen als ursprünglichste der drei Abteilungen angesehen (Panzagl et al. 1996; Löffelhardt et al. 1997).

Grünalgen unterteilt man in Chlorophyta, welche aus Chloro-, Ulvo-, Prasino- und Trebouxiophyceae bestehen, und die Charophyceae, die zusammen mit den Landpflanzen als Streptophyta bezeichnet werden (Bremer 1985; Bremer et al. 1987). Unter den Charophyceen werden die Chlorokybales mit dem einzigen Vertreter *Chlorokybus atmophyticus* als ursprünglichste Gruppe angesehen, während die Frage nach der höchstentwickelten Gattung noch ungeklärt ist (Petersen et al. 2003). Als vielversprechendste Kandidaten werden auch aufgrund molekularer Belege entweder *Coleochaete* oder die morphologisch deutlich komplexere Armleuchteralge *Chara* gehandelt (Graham et al. 2000; Karol et al. 2001; Petersen et al. 2003). Aus den Charophyceen entwickelten sich unter komplexer Umgestaltung von Morphologie und Physiologie die Landpflanzen, welche z.B. durch einen perfektionierten Verdunstungsschutz sukkulenter Vertreter seitdem sogar extrem lebensfeindliche Regionen wie Wüsten erobern konnten. Als primitivste Landpflanzen gelten die Lebermoose (Qui et al. 1998; Wellman und Gray 2000) mit ihrem charakteristischen Vertreter *Marchantia polymorpha*.

Ein ungelöstes Problem war bisher die eindeutige Klassifizierung der einzelligen Grünalge *Mesostigma viride*. Wurde sie zunächst als primitive Prasinophyceae klassifiziert (Mattox und Stewart 1984), favorisierten spätere molekulare Untersuchungen eine Position basal zu allen Grünalgen (Lemieux et al. 2000; Turmel et al. 2002) oder vor den Streptophyten (Melkonian 1989; Bhattacharya 1998; Marin und Melkonian 1999; Martin et al. 2002). In einem kürzlich

1. Einleitung

erschienenen Review-Artikel über die Evolution von Charophyceae und Landpflanzen wurde ihr mit dem Titel „The enigma of *Mesostigma*“ sogar ein eigenes Kapitel gewidmet (McCourt et al. 2004).

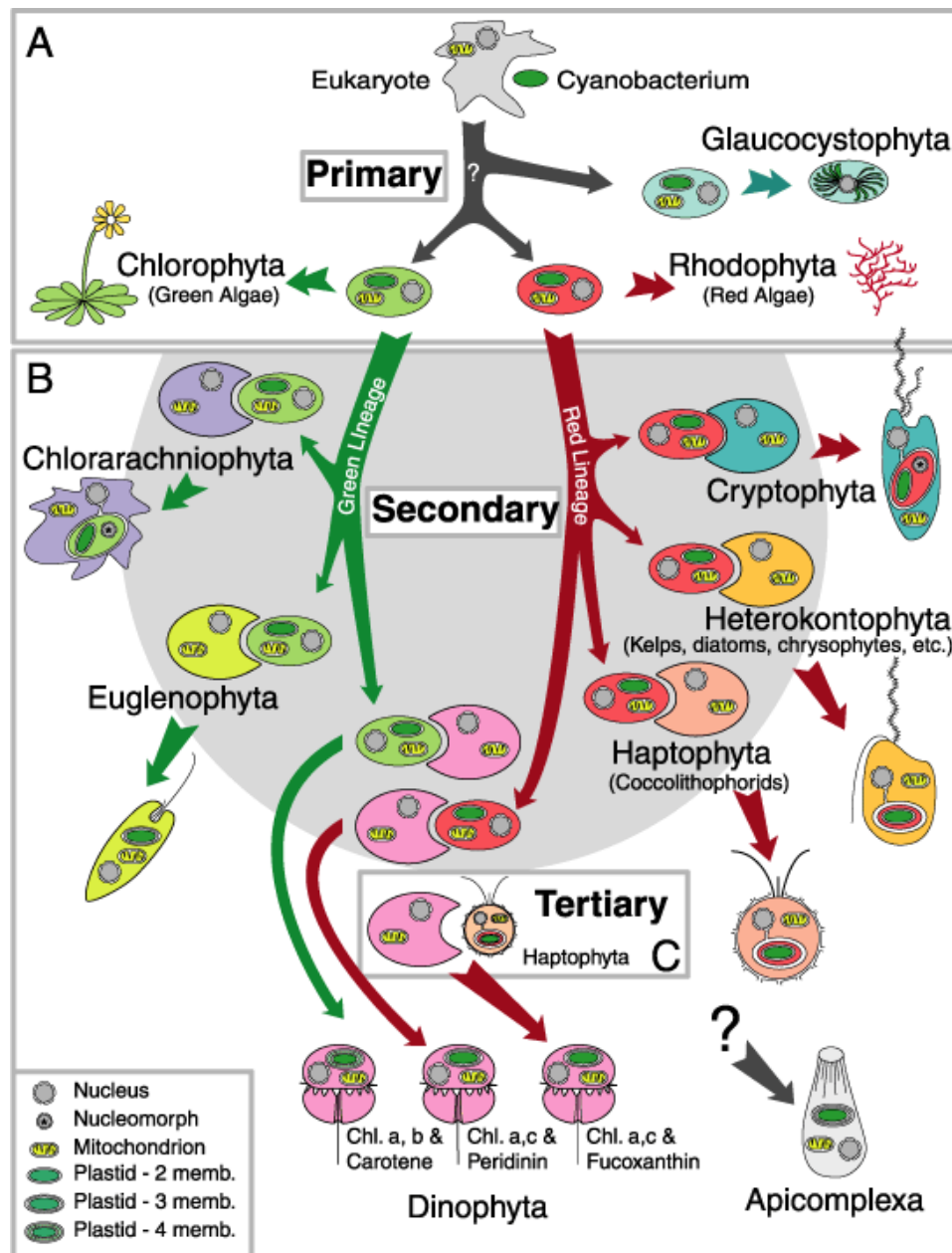


Abbildung 1: Evolution photosynthetischer Eukaryoten

A) Entstehung der drei Hauptlinien der Glaucocystophyta, Grünalgen und Rotalgen durch primäre Endosymbiose mit einem Cyanobakterium. B) Sekundäre Endosymbiosen mit Grün- („Green Lineage“) oder Rotalgen („Red Lineage“). C) Entstehung von Dinophyten mit Fucoxanthin-haltigen komplexen Plastiden durch tertiäre Endosymbiose mit einem Haptophyten. Abbildung modifiziert nach Delwiche 1999.

1. Einleitung

1.2 Die Evolution eukaryotischer Algen – sekundäre und tertiäre Endosymbiosen

„Tiere + Chromatophoren = Pflanzen

...

Diatomeen – Chromatophoren = Tiere“

nach Constantin Mereschkowsky 1910

(Chromatophor = Plastid)

Algen besetzen als Hauptprimärproduzenten entscheidende Schlüsselpositionen im globalen Kohlenstoff- und Stickstoffkreislauf. Der Großteil des eukaryotischen Phytoplanktons der Weltmeere wird jedoch nicht von Grün- oder Rotalgen bestimmt, sondern ist das Ergebnis sekundärer Endosymbiosen, bei denen Grün- oder Rotalgen von ebenfalls eukaryotischen Wirtszellen aufgenommen und zum Plastiden reduziert wurden (siehe Abbildung 1). Die Plastiden dieser im Weiteren als „komplexe Algen“ bezeichneten Abteilungen sind im Gegensatz zu den primären Algenlinien von drei oder vier Membranen umgeben.

Euglenophyta und Chlorarachniophyta haben ihre Plastiden in zwei unabhängigen (sekundären) Endosymbiosen mit Grünalgen erworben (Gibbs 1978; Delwiche 1999; Archibald und Keeling 2002). Die Symbiose mit einer Rotalge führte zur Entstehung der Haptophyta, Cryptophyta, Heterokontophyta (z.B. Braunalgen, Diatomeen) und Peridinin-haltigen Dinoflagellaten (Delwiche 1999; Palmer 2003; siehe Abb. 1). Durch molekularbiologische Untersuchungen konnte zudem eine photosynthetische Vergangenheit der Apicomplexa, die parasitäre Einzeller wie den Malariaerreger *Plasmodium falciparum* enthalten, nachgewiesen werden. Ob ihr heterotropher Plastid (Apicoplast) auf eine Rot- oder Grünalge zurückgeht, wird noch kontrovers diskutiert (Funes et al. 2002; Waller et al. 2003; Funes et al. 2003). Die Analyse nukleärer Gene zeigt eindeutig die Verwandtschaft von Apicomplexa mit Dinoflagellaten und den aplastidären Ciliaten (z.B. dem Pantoffeltierchen *Paramecium*) (Rodríguez-Ezpeleta et al. 2005), weshalb diese zu dem Superensemble Alveolata zusammenfaßt wurden (Van de Peer und De Wachter 1997). Nach wie vor umstritten ist die Anzahl der Endosymbiose-Ereignisse, die zur Entstehung komplexer Algen mit roten Plastiden führten, jedoch deutet die einzigartige Präsenz der plastidären GapC-I, ein Duplikat der cytosolischen GapC, in dieser Algengruppe auf einen gemeinsamen Ursprung der Plastiden hin (Teich 2002). Die sehr weit reichende Chromalveolaten-Hypothese postuliert zudem eine gemeinsame Entstehung der so genannten „Chromista“ (Haptophyta, Cryptophyta und Heterokontophyta) und Alveolata (Dinoflagellaten, Apicomplexa, Ciliaten) durch eine einzige sekundäre Endosymbiose mit einer Rotalge (Cavalier-Smith 1999). Sie ist

1. Einleitung

jedoch äußerst umstritten, da sie auch eine photosynthetische Vergangenheit von Ciliaten postuliert, für die es bisher keinerlei Belege gibt.

1.3 Algenabteilungen mit komplexen Rhodoplasten

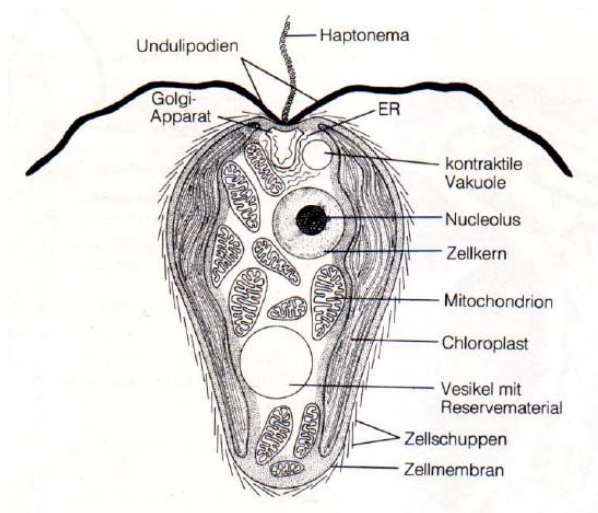
Ein Großteil der heute lebenden Algen ist durch eine sekundäre Endosymbiose entstanden. Komplexe Algen der roten Linie zeichnen sich durch den Besitz des akzessorischen Pigments Chlorophyll c aus. Ihre Morphologie variiert von einzelligen Dinoflagellaten über Kolonien (z.B. der Haptophyt *Phaeocystis*) bis hin zu komplexen mehrzelligen Arten wie den hunderte Meter langen Seetangen (*Macrocystis*) aus der Abteilung der Braunalgen, die zu den größten Organismen der Welt gehören. Erst die mikroskopische Betrachtung enthüllt die extreme Formenvielfalt einzelliger Arten durch externe „Panzer“, die aus Karbonat (Haptophyten) oder Silikat (Diatomeen) bestehen und als fossile Sedimente auf dem Meeresgrund entscheidende Bedeutung für die biogeochemischen Kreisläufe haben.

1.3.1 Haptophyta

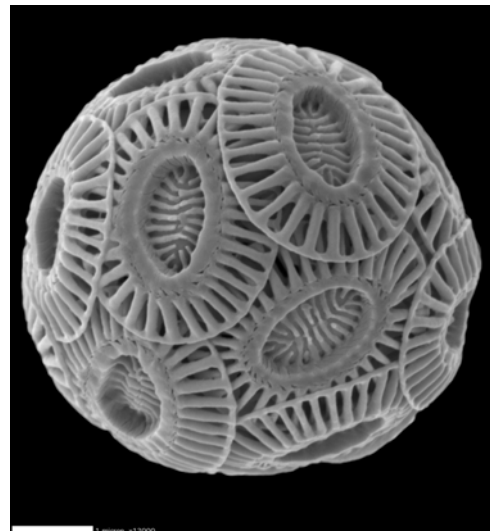
Die Abteilung der Haptophyten setzt sich aus primär marinen, biflagellaten, unizellulären Arten zusammen. Das charakteristische Merkmal dieser Algengruppe ist das namensgebende Haptonem (siehe Abb. 2a), welches für das Anheften an den Untergrund oder das Einfangen von Beute verwendet wird. Man unterteilt Haptophyten in zwei distinkte Klassen: Prymnesiophyceae und Pavlovophyceae. Während die Klasse der Pavlovophyceae nur aus einer Ordnung besteht (Pavlovaales), setzen sich Prymnesiophyceae aus den Prymnesiales, Phaeocystales, Isochrysidales und Coccolithales zusammen (Edwardsen et al. 2000). Eine Besonderheit der Coccolithales (Coccolithophoren) sind die externen Kalkschuppen (siehe Abb. 2b), so genannte Coccolithen, die von mehr als 200 Arten (Graham und Wilcox 2000) synthetisiert werden. Insgesamt bedecken sedimentierte Überreste dieser Arten in etwa die Hälfte des Meeresbodens ($= \frac{1}{3}$ der Erdoberfläche) und sind somit für den größten Langzeit-Niederschlag von anorganischen Kohlenstoffverbindungen verantwortlich. Als Leitfossilien dienen sie den Geologen zur Datierung von Gesteinsschichten und erlauben damit präzise Offshore-Bohrungen zur Erschließung von Erdöllagerstätten. Ihre größte Verbreitung fanden Haptophyten in der Kreidezeit vor 144 bis 65 Mio. Jahren, deren Überreste noch heute als Kreidefelsen auf Rügen und vor Dover weithin sichtbar sind.

1. Einleitung

Rezente Haptophyten beeinflussen die globale Klimaentwicklung, da bei Massenvorkommen von Coccolithophoren die Kalkschuppen das Sonnenlicht in solchem Maße reflektieren, dass hierdurch eine Abkühlung der Erdoberfläche nachgewiesen werden konnte (Graham und Wilcox 2000). Die Arten *Emiliania* (Abb. 2b) und *Phaeocystis* produzieren zudem Dimethylsulfoniopropionat (DMSP) als Osmoregulator bzw. Frostschutz (Graham und Wilcox 2000), welches in der flüchtigen oxidierten Form DMSO als natürlicher Kondensationskeim für die Wolkenbildung dient (Malin und Kirst 1997). Negative Schlagzeilen machen immer wieder toxische Algenblüten von *Prymnesium*-Arten (Abb. 2a), die Massenfischsterben verursachen. Neben dem für komplexe Algen mit Rhodoplasten typischen Chlorophyll c verfügen die meisten Haptophyten über das Pigment Fucoxanthin oder ein Derivat davon (19'Hexanoyloxyfucoxanthin, 19'Butanoyloxyfucoxanthin).



a)



b)

Abbildung 2: Haptophyten

a) Schemazeichnung von *Prymnesium parvum* (aus Margulis und Schwartz 1989); b) EM-Aufnahme von *Emiliania huxleyi*. Der weiße Balken in der linken unteren Ecke repräsentiert 1 µm (zur Verfügung gestellt von Jeremy Young, Natural History Museum London).

1.3.2 Cryptophyta

Aus ökologischer Sicht handelt es sich bei den Cryptophyten (siehe Abbildung 3) um eine eher unspektakuläre Algenabteilung, aufgrund ihrer Morphologie stellen die biflagellaten Einzeller jedoch ein evolutives Highlight dar. Ebenso wie Haptophyten ist ihr Plastid von vier Membranen umgeben, allerdings besitzen Cryptophyten im periplastidären Kompartiment zwischen der zweiten und dritten Plastidenmembran als Relikt der sekundären Endosymbiose

1. Einleitung

den reduzierten Zellkern der aufgenommenen Rotalge. Dieses so genannte Nukleomorph stellt ein „missing link“ bei der Etablierung komplexer Plastiden dar. Das Nukleomorph von *Guillardia theta* (Abb. 3) enthält drei Chromosomen mit einer Genomgröße von etwa 550 kb, die unter anderem für 30 essentielle plastidäre Proteine kodieren (Douglas et al. 2001). Obwohl das Nukleomorph ein extrem dichtgepacktes „Minigenom“ (ungefähr ein Gen/kb) darstellt, enthalten 17 der Protein-kodierenden Gene im 5' Bereich noch typische spleißosomale Introns mit „GT-AG“ Grenzen (Douglas et al. 2001). Cryptomonaden besitzen mit dem Nukleus, dem Nukleomorph, sowie den Plastiden- und Mitochondriengenomen somit insgesamt vier verschiedene Genome. Im Rahmen dieser Arbeit wurden mit *Hanusia phi* und *Rhodomonas salina* zwei weitere Vertreter der Cryptophyten phylogenetisch analysiert.

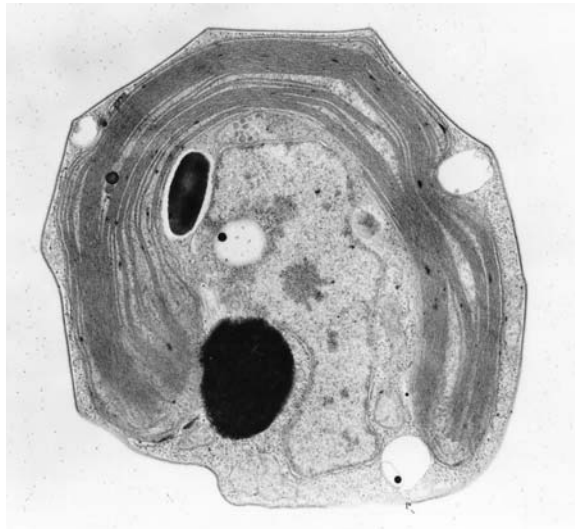


Abbildung 3: *Guillardia theta*

TEM Aufnahme von *Guillardia theta* (Cryptophyta; zur Verfügung gestellt von Sven B. Gould, Philipps-Universität Marburg).

1.3.3 Heterokontophyta (inklusive Diatomeen)

Unter der Bezeichnung Heterokontophyta (Stramenopile) werden Organismen zusammengefaßt, deren Zellen zumindest in einem Lebensstadium über eine heterokonte Begeißelung verfügen. Diese zeichnet sich durch eine lange, nach vorne gerichtete Geißel, die mit Mastigonemen (Flimmerhärchen) besetzt ist und eine kürzere, nach hinten gerichtete Geißel ohne Mastigonemen aus (van den Hoek et al. 1993). Zu den Heterokontophyta zählen sowohl photosynthetische Abteilungen inklusive der Kieselalgen (Diatomeen; siehe Abb. 4), als auch die meist parasitären Oomyceten wie der Erreger der Kraut- und Knollenfäule

1. Einleitung

Phytophthora infestans. Die Algen besitzen eine beeindruckende Variabilität des Phänotyps – von einzelligen Vertretern der Chrysophyceae (Goldalgen) bis hin zu den riesigen Seetangen aus der Klasse der Phaeophyceae (Braunalgen), die vom Erscheinungsbild an Landpflanzen erinnern, in ihrer Morphologie jedoch nichts mit ihnen gemein haben.

Insbesondere die riesige Klasse (mehr als 10.000 Arten) der einzelligen Diatomeen (z.B. *Phaeodactylum*, *Thalassiosira*) stechen mit der extremen Variabilität ihrer Silicat-Wände (siehe Abb. 4) aus den Heterokontophyten hervor. Sie besitzen eine spezialisierte Vakuole, die ihnen über die Speicherung von Nitrat und Phosphat eine größere Toleranz gegenüber Nährstoffschwankungen verschafft (Falkowski et al. 2004). Ihre Transformierbarkeit macht die Diatomee *Phaeodactylum tricornutum* zu einem Modelorganismus der molekularen Zellbiologie von Algen und erlaubt die subzelluläre Lokalisation von Proteinen mit Hilfe von GFP-Reportergenkonstrukten nachzuweisen (Kilian und Kroth 2005). Eine metabolische Besonderheit ist eine zusätzliche mitochondriale Glykolyse, die bisher in keiner anderen Organismengruppe nachgewiesen wurde (Liaud et al. 2000; persönliche Mitteilung Carolina Rio Bartulos).

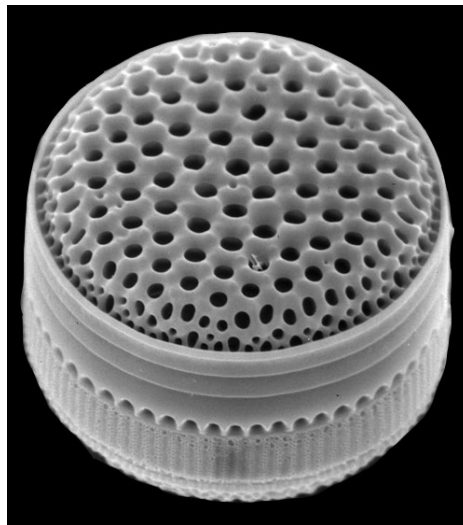


Abbildung 4: Diatomee

EM-Aufnahme einer Diatomee (zur Verfügung gestellt von Jeremy Young, Natural History Museum London).

1.3.4 Dinophyta

Dinophyten (= Dinoflagellaten) sind als Symbionten von Hohltieren wie Seeanemonen und Korallen essentieller Bestandteil der Ökologie von Korallenriffen, können jedoch auch durch toxische Algenblüten („Rote Tiden“) zu massenhaften Sterben von Fischen und Muscheln führen. Als Einzeller faszinieren sie durch die extreme Formenvielfalt ihrer Zellwände aus Dinosporin (siehe Abb. 5). Aber auch ohne Mikroskop bieten beispielsweise *Lingulodinium*

1. Einleitung

polyedrum (Abb. 5) oder *Noctiluca scintillans* durch die Fähigkeit zur Biolumineszenz ein optisches Highlight, das sich als Meeresleuchten beobachten läßt. Ungefähr die Hälfte der Dinoflagellaten-Arten besitzen keinen Plastiden, wenngleich molekulare Analysen darauf hindeuten, dass diese Heterotrophie sekundär durch den Verlust der komplexen Plastiden entstand (Saldarriaga et al. 2001). Der verbreitetste Plastidentyp wird von Peridinin, einem für Dinoflagellaten spezifischen Xanthophyll, dominiert. Allerdings besitzen einige Arten auch Plastiden, die eine abweichende Pigmentierung aufweisen und phylogenetische Untersuchungen belegen, dass diese mit einem unabhängigen Ursprung der Plastiden durch tertiäre Endosymbiosen korrelieren. Besonders deutlich wird dies bei Plastiden mit 19'Hexanoyloxyfucoxanthin, dem typischen Pigment von Haptophyten, aber es wurden auch Diatomeen und Cryptophyten aufgenommen und zum Plastiden reduziert (Tengs et al. 2000; Hacket et al. 2003; Chesnick et al. 1997). Beispiele für unabhängige sekundäre Endosymbiosen mit Prasinophyten liefern die Chlorophyll b haltigen Arten *Lepidodinium viride* und *Gymnodinium chlorophorum* (Watanabe und Sasa 1991; Schnepf und Elbraechter 1999).

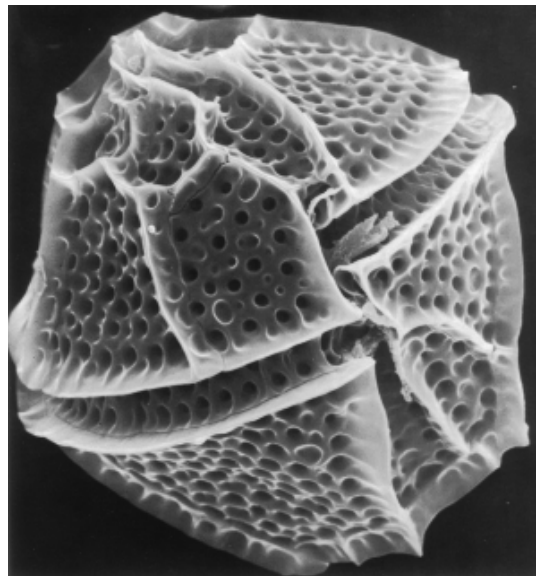


Abbildung 5: *Lingulodinium polyedrum*

EM-Aufnahme von *Lingulodinium polyedrum* (Dinophyta; zur Verfügung gestellt von J. Woodland Hastings, Harvard University Cambridge).

Auch auf genetischer Ebene überraschen Dinoflagellaten durch einige Kuriositäten. Das Plastidengenom ist fragmentiert und besteht nur noch aus so genannten „Minicircles“, die bei *Heterocapsa triquetra* z.T. nur jeweils ein Gen enthalten (Zhang et al. 1999). Unter Eukaryoten einzigartig ist zudem ihre nukleär kodierte Form II Rubisco (Ribulose-1,5-

1. Einleitung

bisphosphat Carboxylase/Oxygenase), die sich nur aus zwei großen Untereinheiten zusammensetzt und zu der aus anaeroben Bakterien homolog ist (Morse et al. 1995, Morden und Sherwood 2002).

Phylogenetische Analysen etwa der 18S rDNA belegen eindeutig den gemeinsamen Ursprung von Dinophyta, nicht-photosynthetischen Apicomplexa (z.B. Malariaerreger) und Ciliaten (z.B. Pantoffeltierchen) weshalb sie im Superensemble Alveolata zusammengefaßt wurden (Van de Peer und De Wachter 1997). Eine kürzlich erschienene Analyse von 143 konkatenierten, nukleär kodierten Proteinen zeigt zudem eine eindeutige Assoziation der Alveolata zusammen mit den Heterokontophyta (Stramenopile; Rodríguez-Ezpeleta et al. 2005).

1.4 Der Primärmetabolismus photosynthetischer Eukaryoten

Die Gewinnung von Energie ist für alle Lebewesen essentiell. Algen und Pflanzen synthetisieren im Rahmen der Photosynthese Zucker, die letztlich auch allen heterotrophen Eukaryoten als Energielieferant dienen. Sowohl der anabole Calvin-Zyklus als auch die katabole Glykolyse sind zentrale Bestandteile des Primärstoffwechsels (siehe Abb. 6).

1.4.1 Photosynthese und Calvin-Zyklus

Bei der Photosynthese wird mittels der absorbierten Lichtenergie an der Thylakoidmembran der Chloroplasten zunächst ein elektrochemischer Elektronengradient aufgebaut und die Energie über eine ATPase auf ATP übertragen. Während des Elektronentransports kommt es zur Spaltung von Wasser (Photolyse), wodurch Sauerstoff freigesetzt wird, der die Grundlage allen aeroben Lebens bildet. Das dabei entstehende Reduktionspotential wird zur Herstellung von NADPH genutzt. Dieses wird zusammen mit ATP für die Fixierung des CO₂ verwendet, sodass die absorbierte Lichtenergie in Form von Glucose gespeichert ist.

Das zentrale, anabole Element der Photosynthese von Cyanobakterien und photosynthetischen Eukaryoten ist der 1956 entdeckte Calvin-Benz-Bassham oder kurz Calvin-Zyklus (= reduktiver Pentosephosphat Zyklus; siehe Abb. 6, links). Er ist der einzige Weg zur CO₂-Fixierung in Eukaryoten und auf das Stroma der Plastiden von Pflanzen und Algen beschränkt. Verschiedene Zwischenprodukte bilden die Grundlage für eine Reihe weiterer essentieller Stoffwechselwege des Plastiden wie der Fettsäure- oder Isoprenoidsynthese (Abb. 6).

1. Einleitung

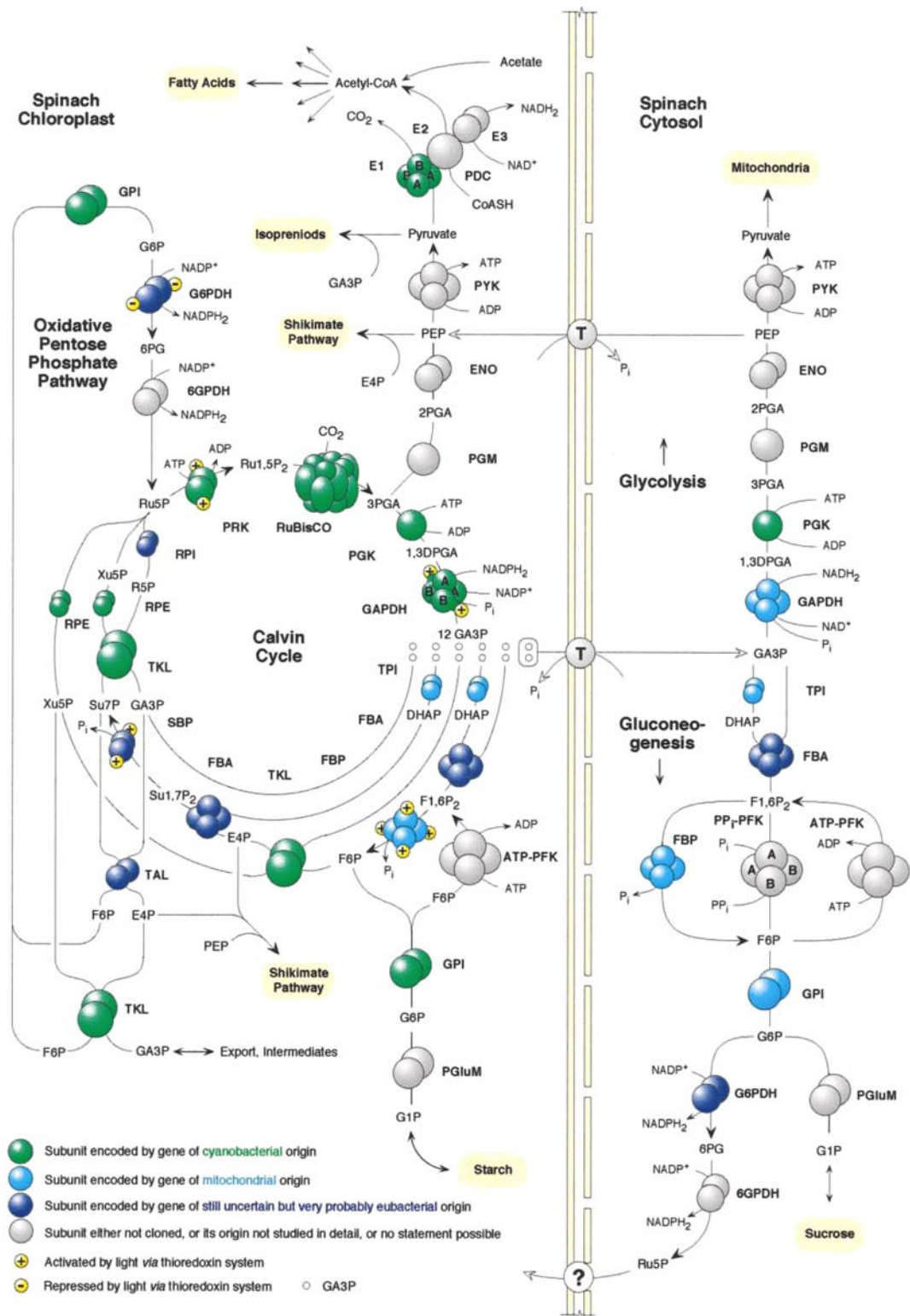


Abbildung 6: Glykolyse und Calvin-Zyklus

Dargestellt ist der zentrale Kohlenhydratstoffwechsel mit den entsprechenden Enzymen am Beispiel des Spinats. FBA, Fructose-1,6-bisphosphat Aldolase; FBP, Fructose-1,6-bisphosphatase; GAPDH, Glycerinaldehyd-3-phosphat Dehydrogenase; PGK, 3-Phosphoglyceratkinase; RPI, Ribulose-5-phosphat Isomerase; PRK, Phosphoribulokinase; RPE, Ribulose-5-phosphat 3-Epimerase; SBP, Sedoheptulose-1,7-bisphosphatase; TKL, Transketolase; TPI, Triosephosphat Isomerase; TAL, Transaldolase; GPI, Glucose-6-phosphat Isomerase; G6PDH, Glucose-6-phosphat Dehydrogenase; 6GPDH, 6-Phosphogluconat Dehydrogenase; PGluM, Phosphoglucomutase; PGM, Phosphoglyceromutase; PFK, Phosphofructokinase (Pyrophosphat und ATP-abhängig); ENO, Enolase; PYK, Pyruvat Kinase; PDC, Pyruvat Dehydrogenase Complex (E1, E2, E3 Komponenten); und T, Translokator. Aus Martin und Herrmann 1998.

1. Einleitung

Der Calvin-Zyklus (Abb. 6, links) setzt sich aus drei Teilen zusammen und besteht aus der Carboxylierungsreaktion (CO_2 -Fixierung), der Reduktion des Kohlenstoffs sowie der Regeneration des CO_2 -Akzeptors Ribulose-1,5-bisphosphat. Während der Carboxylierung wird über die Rubisco (Ribulose-1,5-bisphosphat Carboxylase/Oxygenase) aufgenommenes CO_2 auf Ribulose-1,5-bisphosphat übertragen. Dabei entstehen letztlich zwei Moleküle 3-Phosphoglycerat, die nun unter Verwendung von ATP durch die Phosphoglyceratkinase (PGK) zuerst phosphoryliert und schließlich von der Glycerinaldehyd-3-phosphat Dehydrogenase (GAPDH) mittels NADPH zu Glycerinaldehyd-3-phosphat (GA3P) reduziert werden. Die notwendige Energie in Form von ATP sowie den Reduktionsäquivalenten (NADPH) stammen aus der Lichtreaktion der Photosynthese. Der C_3 -Zucker GA3P wird wieder zur Regeneration des Ribulose-5-phosphat verwendet, wobei nach einer weiteren Phosphorylierungsreaktion durch die Phosphoribulokinase (PRK) wieder das CO_2 -Akzeptormolekül Ribulose-1,5-bisphosphat zur Verfügung steht (Abb. 6). Durch die Fixierung jeweils eines Moleküls CO_2 entsteht nach drei Durchläufen des Calvin-Zyklus ein Molekül GA3P. Dieses wird in Landpflanzen entweder in die osmotisch inaktive Reservestärke umgewandelt oder über den Triosephosphat-Translokator in das Cytosol exportiert, wo es für weitere anabole beziehungsweise katabole Reaktionen zur Verfügung steht (Literatur zu diesem Abschnitt: Voet und Voet 1992; Lüttge et al. 1999).

1.4.2 Glykolyse

Im Gegensatz zum photosynthetischen Calvin-Zyklus ist die Glykolyse ein universeller Stoffwechselweg, der sowohl in Archaea, Eubakterien als auch im Cytosol von Eukaryoten vorkommt (Fothergill-Gilmore und Michels 1993; Martin und Herrmann 1998). Als Glykolyse bezeichnet man den schrittweisen Abbau von dem C_6 -Zucker Glucose zu Pyruvat (= Brenztraubensäure, C_3 ; siehe rechte Seite von Abb. 6, von unten nach oben). Hierbei wird zunächst die Glucose mittels Phosphorylierung aktiviert und durch die Glucosephosphat Isomerase (GPI) in Fructose-6-phosphat umgewandelt. Nach einer weiteren Phosphorylierung durch die Phosphofructokinase (PFK) wird das entstandene Fructose-1,6-bisphosphat über die Fructosebisphosphat Aldolase (FBA) in Glycerinaldehyd-3-phosphat (GA3P) und Dihydroxyacetonphosphat (DHAP) gespalten. Beide C_3 -Verbindungen können mittels der Triosephosphat Isomerase (TPI) ineinander isomerisiert werden. Für die weitere Phosphorylierung und Oxidation zu 1,3-Diphosphoglycerat (1,3DPGA) mittels der Glycerinaldehyd-3-phosphat Dehydrogenase (GAPDH) dient jedoch ausschließlich GA3P als

1. Einleitung

Substrat. Nach der anschließenden Dephosphorylierung des 1,3DPGA durch die Phosphoglyceratkinase (PGK) wird 3-Phosphoglycerat (3PGA) sukzessive zu Pyruvat umgewandelt (siehe Abb. 6).

Nach anfänglichem Einsatz von Energie wird letztlich in der Glykolyse aus der Glucose nur eine geringe Menge ATP und NADH (Reduktionsäquivalent) gewonnen (siehe Abb. 6, rechts). Der Großteil der Energie ist weiterhin im Pyruvat gespeichert und wird nach weiterem Abbau im mitochondrialen Zitronensäurezyklus und der Endoxidation in Form von ATP freigesetzt (Literatur zu diesem Abschnitt: Voet und Voet 1992; Lüttge et al. 1999).

1.5 Ursprung und Evolution der Gene für Glykolyse und Calvin-Zyklus

Cytosolische Glykolyse und plastidärer Calvin-Zyklus beinhalten homologe Isoenzyme, die kompartimentspezifisch die jeweils gegenläufige katabole oder anabole Reaktion katalysieren. Insgesamt besitzen beide Reaktionswege fünf solcher Isoenzympaare (PGK, GAPDH, TPI, FBA und FBP).

Glykolyse und die anabole Gluconeogenese sind universeller Bestandteil des Metabolismus von Archaea, Eubacteria und Eukarya. Während sich in Eukaryoten die Gene für die Zellstruktur und das genetische Transkriptions- und Translationsystem von Archaeabakterien ableiten lassen, hat der eukaryotische Primärmetabolismus einen eubakteriellen Ursprung (Martin et al. 1993; Martin und Schnarrenberger 1997; Ronimus und Morgan 2003). Im Laufe der Evolution wurde die Glykolyse aus den α -Proteobakterien (Mitochondrien) dementsprechend in das Cytosol verlagert, wohingegen die Endoxidation für die ATP-Gewinnung in den Mitochondrien verblieb (s.o.).

Eukaryoten haben die Photosynthese im Zuge der primären plastidären Endosymbiose von Cyanobakterien „geerbt“. Heutige Plastidengenome kodieren nur noch für ca. 20 bis 200 Proteine während das Cyanobakterium *Synechocystis* 3168 verschiedene Proteine besitzt. Für die Aufrechterhaltung der vollen Funktionalität der Plastiden werden jedoch schätzungsweise ungefähr 2000 Proteine benötigt (Leister und Schneider 2004). Dieses Defizit wird durch einen massiven Import nukleär kodierter Proteine in die Organellen ausgeglichen. Im Rahmen der Endosymbiosen, die zur Entstehung von Plastiden (/Mitochondrien) führten, erfolgte ein massiver Transfer genetischen Materials in den Nukleus der Wirtszelle (Leister und Schneider 2004). Kurzzeitig existierten zwei Kopien des gleichen Gens im Organellen- und im Kerngenom. Erst nach der Anpassung der nukleären Kopien an das genetische System der

1. Einleitung

eukaryotischen Wirtszelle, erfolgreicher Expression und dem Erwerb einer Präsequenz, welche den Reimport des Proteins in den Plastiden (/das Mitochondrium) gewährleistet, konnte die plastidäre (/mitochondriale) Kopie funktionell ersetzt werden und ging verloren (Martin und Herrmann 1998; siehe Brennicke et al. 1993 für Gentransfer aus Mitochondrien). Durch den Transfer der Gene aus den Organellen in den Nukleus der Wirtszelle erfolgte neben der zellulären und biochemischen Symbiose nun auch eine irreversible genetische Verzahnung von Endosymbiont und Wirt.

1981 stellte Weeden die Hypothese auf, dass das Produkt eines transferierten Gens wieder in das ursprüngliche Organell importiert wird („product specificity corollary“, Weeden 1981). Sie postuliert, dass nach der plastidären Endosymbiose alle plastidären Proteine cyanobakteriellen Ursprungs sind und die cytosolischen Enzyme die Wirtszelle repräsentieren. Anhand phylogenetischer Analysen der Enzyme für Glykolyse und Calvin-Zyklus wurde jedoch für Landpflanzen am Beispiel des Spinat gezeigt, dass diese Grundannahme zu stark vereinfacht ist und nur für einen Teil der Gene zutrifft (Martin und Schnarrenberger 1997). Ein Teil der nukleär kodierten plastidären Enzyme geht auf duplizierte Gene der Wirtszelle zurück, die ein Transitpeptid erworben haben. Der Primärstoffwechsel zeigt damit insgesamt eine mosaikartige Zusammensetzung von Enzymen cyano- und proteobakteriellen Ursprungs (Martin und Schnarrenberger 1997, Martin und Hermann 1998). Allerdings bestimmt der Zufall, welches Gen im Nukleus etabliert wird, sodass sich das Schicksal individueller Gene nach Endosymbiosen nicht prognostizieren lässt.

1.6 Spleißosomale Introns – Zeugnisse eukaryotischer Evolution

Eine überraschende Entdeckung der siebziger Jahre des letzten Jahrhunderts war, dass in eukaryotischen Genen kodierende Bereiche (Exons) durch nichtkodierende Bereiche (Introns) unterbrochen sind. Der größte Anteil dieser Introns wird von konservierten Konsensussequenzen flankiert, wobei ein Intron am 5' Ende meist mit „GT“ beginnt und am 3' Ende mit „AG“ endet (Tolstrup et al. 1997). Das Herausschneiden der nichtkodierenden Sequenzen, das so genannte Spleißen, erfolgt auf mRNA Ebene und wird vom Spleißosom, einem Komplex bestehend aus mehreren snRNPs (small nuclear ribonucleoprotein), die sich wiederum aus RNAs und Proteinen zusammensetzen, vermittelt. Bei der Spleißreaktion wird das Intron in Form eines „Lassos“ (Lariat-Struktur) präzise herausgeschnitten und dabei das 3' Ende des ersten und das 5' Ende des nachfolgenden Exons miteinander verknüpft (Jurica

1. Einleitung

und Moore 2003). Mittlerweile wurden auch in einigen Prokaryoten so genannte selbstspleißende Introns nachgewiesen, die autokatalytisch und nicht durch ein Spleißosom entfernt werden (Robart und Zimmerly 2005).

Da eine Aminosäure von drei Nukleotiden kodiert wird (Triplett-Code) kann sich ein Intron (auf DNA-Ebene) in drei so genannten Phasen befinden. Als Phase „0“ bezeichnet man dabei die Insertion des Introns vor dem ersten Nukleotid, Phase „1“ nach dem ersten und Phase „2“ nach dem zweiten Nukleotid der abgeleiteten Aminosäure (siehe Abb. 7). Dies ermöglicht die eindeutige Bestimmung einer Intronposition im Bezug auf die abgeleitete Aminosäureposition und erlaubt basierend auf entsprechenden Alignments den Vergleich homologer Introns über die Artgrenzen hinweg.

16	17	18	19	Position
T	V	P	R	AS-Sequenz
ACG	GTA	CCT	CGA	nt-Sequenz
		↑↑↑		
		Phase 0 1 2		

Abbildung 7: Intronphasen

Dargestellt ist eine beispielhafte Sequenz aus vier Aminosäuren. Pfeile markieren die Intronpositionen 18-0 (Position 18 Phase 0), 18-1 bzw. 18-2.

Obwohl meist als nutzlos („junk“) bezeichnet, wird eine mögliche Bedeutung von Introns im Rahmen der Evolution diskutiert. In einem als „exon-shuffling“ bezeichneten Prozess werden einzelne Module (Exons) kombiniert, wodurch neue Gene entstehen (Long und Cerff 2003; Roy und Gilbert 2006) oder neu erworbene Gene eine Präsequenz aquirieren (Long et al. 1996). Darüber hinaus wurden in Introns regulatorische Elemente für die Expression der Gene entdeckt (Sakamoto et al. 2001, Rombauts et al. 2003, Kim et al. 2004). Im Allgemeinen unterliegen Introns aber im Gegensatz zu kodierenden Bereichen (Exons) keinem selektiven Druck und sammeln so in einer hohen Rate Sequenzmutationen an. Selbst in evolutiv entfernten Organismen wurden jedoch identische Intronpositionen beobachtet (Kersanach et al. 1994). Solche konservierten Intronpositionen können ebenso wie Indels (Sequenzinsertionen und -deletionen) als phylogenetische Marker verwendet werden, deren Erwerb oder Verlust neben phylogenetischen Analysen Aufschluß über die evolutive Vergangenheit gibt.

1. Einleitung

1.7 Ziele der Arbeit

Die vorliegende Arbeit gliedert sich in zwei thematische Schwerpunkte. Der erste Teil (A) befaßt sich auf der Grundlage des GAPDH Systems mit dem Ursprung der Landpflanzen. Die *GapA/B* Genduplikation (Teil A1) korreliert mit dem Übergang vom Wasser zum Landleben und die Präsenz der neu entstandenen *GapB* könnte dazu beitragen, die bisher ungeklärte Position der einzelligen Grünalge *Mesostigma viride* zu enthüllen. Es sollen entsprechende GAPDH Sequenzen von *Mesostigma* etabliert und phylogenetisch analysiert werden, um ihre Stellung zu den Charophyceen – aus denen die Landpflanzen hervorgegangen sind – zu bestimmen.

Neben der *GapA/B* Genduplikation korreliert auch eine zweite Duplikation (*GapC/Cp*) mit dem Übergang zu den Landpflanzen. Ein weiterer Schwerpunkt (A2) bildet der Vergleich der Genstruktur von *GapA*, *GapB*, *GapC* und *GapCp* Genen. Die Positionen von Introns können über lange evolutive Zeiträume konserviert sein und somit die Schlußfolgerungen aus phylogenetischen Analysen bestätigen. In der vorliegenden Arbeit soll das Intronmuster von Grünalgen und Landpflanzen vor und nach diesen Duplikationen untersucht werden, um die Evolution der Genstruktur nachzuvollziehen.

Im zweiten Teil (B) werden die phylogenetischen Beziehungen der eukaryotischen Algenabteilungen anhand von Genen des Primärstoffwechsels untersucht. Der Schwerpunkt liegt hierbei auf komplexen Algen mit Rhodoplasten, zu denen die Haptophyten, Cryptophyten, Diatomeen und Dinophyten zählen. Da für letztere Gruppe auf Wirtszellebene der gemeinsame Ursprung zusammen mit den Apicomplexa (z.B. Malariaerreger) und Ciliaten (z.B. Pantoffeltierchen) nachgewiesen ist (Alveolata), werden auch diese, soweit verfügbar, in die Analysen mit einbezogen. Ziel dieser Arbeit ist es, Rückschlüsse auf den ungeklärten und kontrovers diskutierten Ursprung der Algenabteilungen mit komplexen roten Plastiden zu erhalten.

2. Ergebnisse Teil A - Die Evolution der Streptophyta

Aufgrund morphologischer, ultrastruktureller, biochemischer und molekularer Daten ist bekannt, dass sich Landpflanzen aus den zu den Grünalgen gehörenden Charophyceae entwickelt haben. Innerhalb der Charophyceae sind die genauen evolutiven Stellungen der einzelnen Gattungen jedoch bis heute umstritten. Darüber hinaus wurde zu Beginn dieser Arbeit insbesondere die Position von *Mesostigma viride* in der primären grünen Linie kontrovers diskutiert. Der erste Teil der vorliegenden Arbeit fokussiert auf das GAPDH-System von Grünalgen und terrestrischen Pflanzen. Bereits in einer vorangegangenen Dissertation (Petersen 1999) konnte nachgewiesen werden, dass zwei für Streptophyta entscheidende Duplikation nukleärer Gene der plastidären GapA und cytosolischen GapC mit diesem Übergang zu den Landpflanzen korrelieren. Ziel dieser Arbeit war es nun, die bisher ungeklärte Stellung der Grünalge *Mesostigma* zu den Streptophyta zu bestimmen und die Evolution der GAPDH-Genstruktur im Kontext der Entstehung der Landpflanzen zu untersuchen.

Glycerinaldehyd-3-phosphat Dehydrogenasen katalysieren als Tetramer die reversible oxidative Phosphorylierung von Glycerinaldehyd-3-phosphat zu 1,3-Diphosphoglycerat (Cerff 1979; Cerff und Chambers 1979; Cerff 1982). Grün- und Rotalgen besitzen eine plastidäre GapA cyanobakteriellen und eine cytosolische GapC proteobakteriellen (mitochondrialen) Ursprungs. Daneben enthalten die Plastiden der Landpflanzen eine GapB, die durch eine Genduplikation der *GapA* entstanden ist (Brinkmann et al. 1989). Die GapB erwarb zudem eine C-terminale Extension (CTE), die eine Regulation des Enzyms über das Thioredoxin-System ermöglicht (Sparla et al. 2002). Durch eine zweite Genduplikation entstand aus der cytosolischen GapC die GapCp. Sie ist Bestandteil der plastidären Glykolyse in heterotrophen Geweben von Landpflanzen (Petersen et al. 2003).

2. Ergebnisse A

2.1 Etablierung und Analyse plastidärer GAPDH-Sequenzen

2.1.1 Etablierung plastidärer GAPDH-Sequenzen mittels degenerierter Primer und der RACE-Technik

Ein von Burkhard Becker (Köln) initiiertes EST-Projekt von *Mesostigma viride* (Simon et al. 2006) bildete die Basis für eine Kooperation, in der mehrere GAPDH-Klone und Algenmaterial der einzelligen Grünalge zur Verfügung gestellt wurden. Die komplette Sequenzierung ausgewählter EST-Klone zeigte, dass diese GapA, GapB und GapC Sequenzen repräsentierten. Bei der GapB und GapC handelte es sich bereits um vollständige cDNA Klone, der längste GapA-Klon war allerdings nur partiell (ab Aminosäureposition 26). Ein in GAPDH-Sequenzen nahe des N-Terminus gelegenes hochkonserviertes Proteinmotiv (GINGFGRI) wurde für die Erstellung eines degenerierten Primers (P054; siehe Kapitel 7.13) verwendet. Die bekannten Sequenzinformationen der GapA dienten darüber hinaus der Erstellung von zwei spezifischen Primern in Minus-Richtung (P462, P463). Ausgehend von isolierter RNA wurden die spezifischen sowie der degenerierte GAPDH Primer für eine semi-nested RT-PCR verwendet, womit die fehlende Sequenz ab Aminosäure 4 etabliert werden konnte (entspricht nt 1-72 von DQ270265).

In einer früheren Arbeit (Petersen 1999) konnte eine GapB cDNA-Sequenz des Lebermooses *Marchantia polymorpha* etabliert werden, die am 3' Ende allerdings partiell ist. Da dieser Bereich eine Extension enthält, die spezifisch für GapB-Sequenzen ist, wurde er im Rahmen dieser Arbeit mittels der 3'RACE-Technik ab AA 344 (nt 1374-1770 von AJ246027) vervollständigt.

Abbildung 8 zeigt einen Vergleich der C-terminalen Extension (CTE) von GapB-Sequenzen der Grünalgen, dem Lebermoos und Samenpflanzen zusammen mit GapA-Sequenzen und CP12. Letzteres ist für die lichtregulierte Komplexbildung und somit Inaktivierung von GAPDH und Phosphoribulokinase verantwortlich (Pohlmeyer et al. 1996; Wedel et al. 1997) und erlaubt somit in der Nacht den Wechsel vom Calvin-Zyklus zum oxidativen Pentosephosphat Weg (Michels et al. 2005). Die GapB rekrutierte von CP12 das GAPDH-bindende Modul, welches nun als CTE (Pohlmeyer et al. 1996) an der Regulation des Zuckerphosphat-Metabolismus in Grünalgen und Landpflanzen beteiligt ist.

2. Ergebnisse A

CP12	100	120	130	
<i>Arabidopsis thaliana</i>	WDEVEELSAAASHARDKKKADGS--	DPLEEYCKDNPETNE	CRTYDN*	
<i>Oryza sativa</i>L.ET-.-A.....D.....*	
<i>Pinus taeda</i>	..T.....MR..M.T.D..K----F.....D.....ED*	
<i>Physcomitrella patens</i>T.E.A.Q..RE.TS-G--K..A....SD.....ED*	
<i>Marchantia polymorpha</i>	...T.....YADKKR.A.-----T..A.....D.....V.ED*	C
<i>Mesostigma viride</i>DKKK.E.EIAKK-KF.E.....D.....V.ED*	P
<i>Chl. reinhardtii</i>	..T.....V..KK.AV...VTLT...AF...A.DAD...V.ED*	1
<i>Cyanophora paradoxa</i>E.EL..QKAQP.-----Q.F..E.....D.....L.ED*	2
<i>C. merolae</i>	..F....E.ER..QAAN.QSE----K..NEV..AD...V.ED*	
<i>Porphyra yezoensis</i>	..A....E.E...QKNAE.K-----E.V..AD.....ED*	
<i>Synechocystis</i> sp.	..A....E.E.A.Q.QQHPTQ----	TT..KF.DE..DAA...I..D*	
<i>S. elongatus</i>	..A....Q.E.A.Q.QQ.SETE---	PFFGD...SE..DAA...LI..D*	
GAPDH	330	340	350	360
<i>Arabidopsis thaliana</i>	RVVDLAHLVASKWPG---AEAVGSGD	PLED	FKTNPADEEC	CKVYD*
<i>Oryza sativa</i>---.AVQ.....D..ETD.....EN*	G
<i>Pinus taeda</i>A.....-.GKG.....D...TD..R..	EA*	a
Marchantia polymorphaAQ.....-V.KP.....E..V...DTD..	RLH.V*	p
<i>C. atmophyticus</i>I..NN...QKADSEEE.....	DKY..A...EAD..	R...E*	B
<i>Mesostigma viride</i>DI..NN...-MVQETGS....	T..N...DAD..	R...ND*	
<i>Arabidopsis thaliana</i>DI..NN.K*			
<i>Oryza sativa</i>DI..NQ.K*			
<i>Pinus sylvestris</i>DI..NN.KS*			G
<i>Marchantia polymorpha</i>N...E..T.*			a
<i>C. atmophyticus</i>S.I..NNWNSTVEQAKELSAV*			p
<i>Mesostigma viride</i>IM.KEMK*			A
<i>Chl. reinhardtii</i>EVT.KK.VA*			

Abbildung 8: Vergleich GapB C-Terminus und CP12

Die reife Untereinheit von CP12 wurde manuell mit der C-terminalen Extension von GapB-Sequenzen aligniert. Konservierte Aminosäurepositionen sind mit einem Punkt (".") und Deletionen mit einem Strich ("-") dargestellt. Die Nummerierung von CP12 (oben) und GAPDH (unten) basiert auf den Protein Sequenzen von *Arabidopsis thaliana* (AAM45071) und *Geobacillus stearothermophilus* (P00362) als Referenz. Zwei universell konservierte Cystein Reste sind dunkelgrau unterlegt. Hellgrau markiert sind Aminosäurepositionen, die zwischen CP12 und GapB konserviert (ähnlich) sind, wobei Aminosäuren mit einem Wert von ≥ 2 in der PAM250 Substitutions-Matrix als „ähnlich“ angesehen werden. *Chl. reinhardtii*, *Chlamydomonas reinhardtii*; *C. merolae*, *Cyanidioschyzon merolae*; *S. elongatus*, *Synechococcus elongatus*; *C. atmophyticus*, *Chlorokybus atmophyticus*.

2.1.2 Phylogenetische Analyse plastidärer GAPDH-Sequenzen

Abbildung 9 zeigt einen MrBayes Consensus Baum basierend auf einer Analyse von 73 GAPDH-Sequenzen, darunter 57 eukaryotischen GapA und GapB-Sequenzen, sowie 16 cyanobakterielle gap2-Sequenzen, die als Außengruppe verwendet wurden. Der Datensatz enthält Streptophyten-Sequenzen von verschiedenen Landpflanzen inklusive der Laubmoose *Physcomitrella* und *Sphagnum* sowie des Lebermooses *Marchantia*. Weiterhin sind Vertreter aller fünf Ordnungen der Charophyceen Coleochaetales (*Coleochaete*), Charales (*Chara*), Zygnematales (*Spirogyra*), Klebsormidiales (*Klebsormidium*) und Chlorokybales

2. Ergebnisse A

(*Chlorokybus*) sowie die einzellige Grünalge *Mesostigma viride* enthalten. Vertreter aller Klassen der Chlorophyta inklusive der Chloro- (*Chlamydomonas*, *Spermatozopsis*, *Scenedesmus*), Ulvo- (*Cladophora*, *Acetabularia*), Trebouxio- (*Chlorella*) und Prasinophyceae (*Scherffelia*) wurden ebenso eingeschlossen. Die beiden verbleibenden primär photosynthetischen Linien sind durch Sequenzen von *Cyanophora* (Glaucocystophyta) und fünf Rotalgen vertreten.

Ein gemeinsamer Ursprung von *GapA* und *GapB*-Genen von Viridiplantae (Streptophyta und Chlorophyta), Rhodophyta und Glaucocystophyta, welche alle durch eine primäre Endosymbiose mit einem Cyanobakterium entstanden sind, wird solide von Bootstrap-Werten und „Bayesian Posterior Probabilities“ zwischen 99 und 100% unterstützt (siehe Abb. 9). Rotalgen zweigen im eukaryotischen Ast basal ab und *Cyanophora* gruppiert zusammen mit der grünen Linie, welche in drei distinkte Substammbäume unterteilt ist (Chlorophyta *GapA*, Streptophyta *GapA*, Streptophyta *GapB*).

Beide *GapA*-Substammbäume (Chlorophyta bzw. Streptophyta *GapA*) sind statistisch nur moderat unterstützt, was den hohen Grad an Sequenzidentität innerhalb der grünen Linie widerspiegelt (75%, nicht gezeigt). Der *GapB*-Substammbaum wird im Gegensatz dazu solide unterstützt (89-100%) und bildet nach der Genduplikation aber vor dem Abspalten von *Mesostigma* einen gemeinsamen Ast, der zweimal länger ist, als für die *GapA*, was auf eine kurze Phase mit geringem Selektionsdruck hindeutet. Der Vergleich der durchschnittlichen Astlängen der Substammbäume nach dem Abzweigen von *Mesostigma* zeigt, dass *GapB*-Gene einen sehr hohen Grad an Sequenzkonservierung besitzen (nicht gezeigt), welche den der entsprechenden *GapA*-Sequenzen übersteigt. Dies deutet darauf hin, dass die neue *GapB* einem massiven Selektionsdruck unterliegt.

Chlorokybus nimmt im langsam evolvierenden *GapB*-Substammbaum (Abb. 9) die basale Position der Charophyceen ein. Die Stellung von *Klebsormidium* (Charophyceae) inmitten der Landpflanzen des *GapA*-Astes erhält keinen signifikanten statistischen Support. Möglicherweise handelt es sich um ein so genanntes „long branch attraction“ Artefakt (LBA; Philippe et al. 2000; Philippe und Germot 2000; Brinkmann et al. 2005), welches durch die schnell evolvierenden Samenpflanzen verursacht werden könnte. Da die statistische Unterstützung für die Gruppierung von Charophyceen und Moosen sowohl im *GapA* als auch im *GapB* Substammbaum gering ist, lassen die Analysen plastidärer GAPDH-Sequenzen keine Rückschlüsse über den nächsten Verwandten der Landpflanzen auf der Seite der Charophyceae zu.

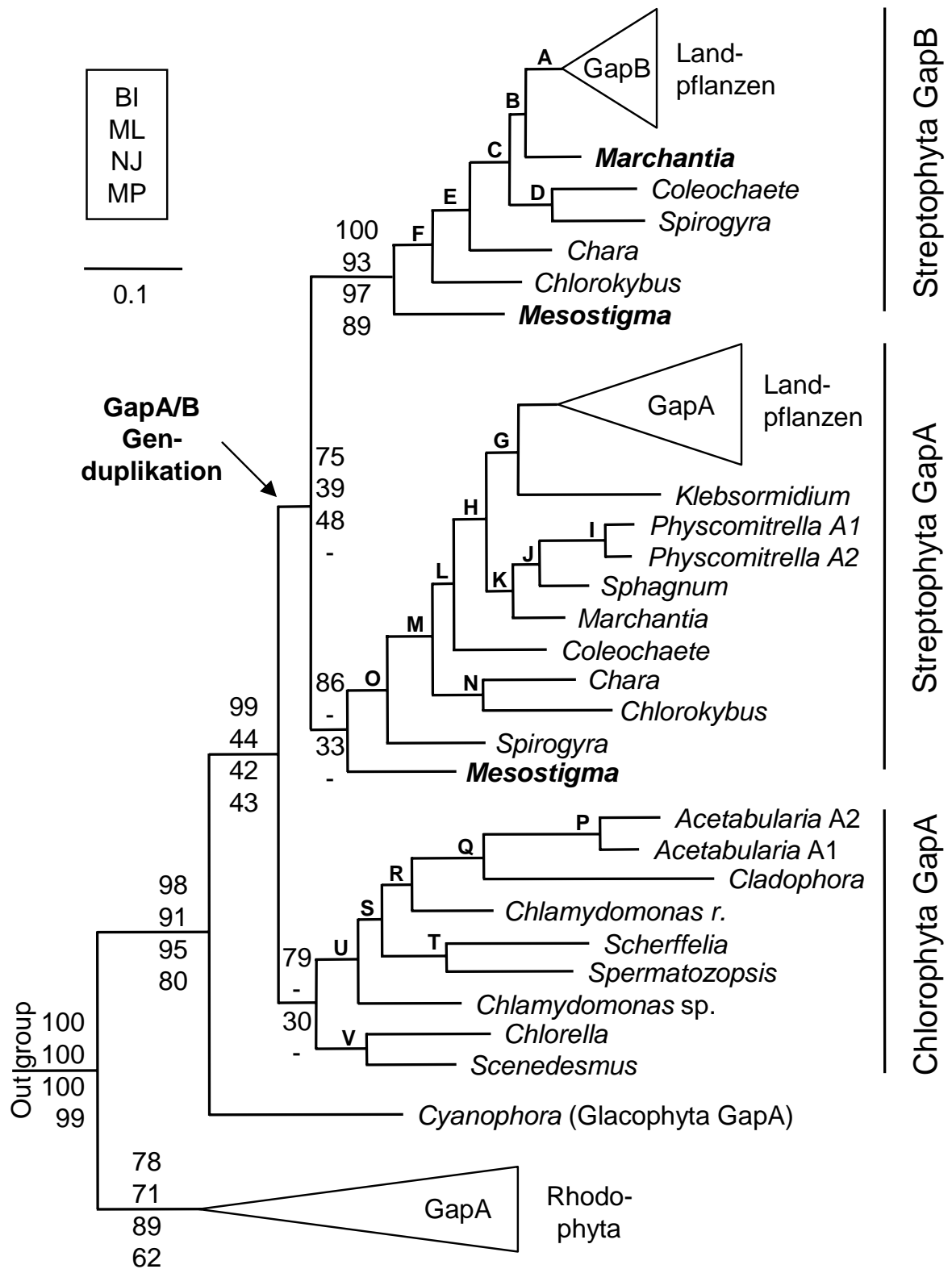
2. Ergebnisse A

Die Präsenz der GapB in *Mesostigma* belegt zusammen mit der basalen Stellung im GapB und GapA Substammbaum, dass diese Grünalge eindeutig mit Streptophyten assoziiert ist und die *GapA/B*-Genduplikation in einem gemeinsamen Vorfahren aller Streptophyta erfolgte. Die Tatsache, dass Chlorophyta-Sequenzen basal zu den Streptophyta gruppieren, deutet darauf hin, dass die Genduplikation nach der Aufspaltung von Chloro- und Streptophyta erfolgte, auch wenn die statistische Unterstützung hierfür niedrig ist.

Abbildung 9: Phylogenetische Analyse von GapA und GapB

MrBayes Consensus Baum (WAG + F + Γ 8 + I Modell) basierend auf 74 GAPDH-Sequenzen und 325 Aminosäurepositionen. Sequenzen, die in dieser Arbeit etabliert wurden, sind **fett** dargestellt. Die horizontale Länge der Dreiecke ist äquivalent zur durchschnittlichen Astlänge. Statistische Support-Werte wurden durch Bootstrap-Analysen oder „Posterior Probabilities“ bestimmt. Support Werte >30% und Posterior Probabilities >70% sind gezeigt. BI, Bayesian Inference; ML, Maximum-Likelihood; NJ, Neighbour Joining Distanz; MP, Maximum Parsimony. Buchstaben an den Verzweigungen stehen für folgende statistische Support Werte (BI|ML|NJ|MP): (A) 100|55|-|40, (B) 73|64|71|40, (C) 100|42|39|44, (D) 100|71|66|67, (E) 97|33|-|-, (F) 94|39|34|-, (G) -|-|-|-, (H) 80|-|-|-, (I) 100|99|100|100, (J) 91|61|51|45, (K) 91|50|54|32, (L) -|-|-|-, (M) 82|31|-|-, (N) 98|44|51|-, (O) 99|49|-|-, (P) 100|100|100|100, (Q) 74|57|47|60, (R) -|-|-|-, (S) -|-|-|38, (T) 99|48|36|57, (U) 88|-|-|-, (V) 99|64|81|46.

2. Ergebnisse A



2. Ergebnisse A

2.2 Die Struktur von GAPDH-Genen

Ausgehend von den phylogenetischen Analysen plastidärer GAPDH-Sequenzen (siehe Abb. 9) wurden die Intronpositionen von GAPDH-Genen der Grünalgen und Landpflanzen untersucht. Das Interesse galt hierbei dem Maß an Konservierung der Genstruktur seit der Entstehung der Streptophyta und insbesondere der basal zu ihnen stehenden Grünalge *Mesostigma* (siehe Kapitel 2.1.2). Neben den *GapA* und *GapB*-Sequenzen aus Streptophyta wurden auch *GapC* und *GapCp*-Sequenzen analysiert.

Zunächst wurden alle verfügbaren genomischen Sequenzen von Landpflanzen sowie den Grünalgen *Chlamydomonas reinhardtii* und *Volvox carteri* durch intensive Datenbankrecherche etabliert und deren Intronpositionen bestimmt. Unter Berücksichtigung dieser Introns wurden anhand vorhandener und in dieser Arbeit etablierter cDNA-Sequenzen der Streptophyta spezifische Primer konzipiert. Mittels PCR-Experimenten mit genomischer DNA amplifizierte Gene wurden kloniert, sequenziert und die Intronpositionen und Phase (siehe Abb. 7) durch den Vergleich mit der entsprechenden cDNA-Sequenz bestimmt. Als Referenz für die Nummerierung der Aminosäurepositionen diente aufgrund historischer Gründe (erste GAPDH-Proteinstruktur) und der Vergleichbarkeit mit älteren Arbeiten (z.B. Kersanach et al. 1994) die *gap1* von *Geobacillus stearothermophilus* (Biesecker et al. 1977; P00362). Es wurden so die Introns in GAPDHs von dem Lebermoos *Marchantia polymorpha*, Charophyceen, dem primitiven Streptophyten *Mesostigma viride* und der Ulvophyceae *Cladophora rupestris* (Chlorophyta) untersucht. Alle neu etablierten Intronpositionen wurden im Vergleich mit den aus Datenbankrecherchen erhaltenen Referenzsequenzen in zwei Matrizes zusammengestellt (Abb. 10 und Abb. 11).

2.2.1 GapA/B (EC-Nummer 1.2.1.13)

Aufgrund der phylogenetischen Analyse (siehe Kapitel 2.1.2) konnte belegt werden, dass die *GapA/B*-Genduplikation in einem gemeinsamen Vorfahren aller Streptophyta inklusive der Grünalge *Mesostigma viride* erfolgte. In der vorliegenden Arbeit wurden genomische *GapA* und *GapB*-Sequenzen der primitiven Landpflanze *Marchantia polymorpha*, der Charophyceen *Chara vulgaris* und *Chlorokybus atmophyticus*, die von *Mesostigma viride*, sowie die genomische *GapA* der Grünalge *Cladophora rupestris* etabliert (Accession Numbers siehe 9.1). Da für die Erstellung der Primer z.T. nur partielle RT-PCR-Klone zur Verfügung standen, konnten auch die genomischen Sequenzen in diesen Fällen nicht vollständig

2. Ergebnisse A

amplifiziert werden. Die fehlenden Bereiche sind in der Intronmatrix (Abb. 10) mit „?“ gekennzeichnet. „TP1“ und „TP2“ (Abb. 10) markieren wahrscheinlich homologe Introns im Transitpeptid der abgeleiteten Aminosäuresequenz von *GapA* und *GapB*-Genen, deren exakte Nummerierung aufgrund geringer Sequenzkonservierung allerdings nicht möglich ist.

Das 1496 bp lange genomische Amplifikat der *GapA* von *Marchantia* enthält drei Introns an den Positionen TP2, 85-2 und 166-1. Die *GapA* von *Chara* und *Mesostigma* enthalten jeweils nur ein Intron (85-2), *Chlorokybus* vier Intronpositionen (116-0, 183-0, 310-2, 333-2) und die Ulvophyceae *Cladophora* besitzt kein Intron.

GapB-Gene enthalten deutlich mehr Introns. So konnten in *Marchantia* sieben Introns nachgewiesen werden (TP1, TP2, 40-0, 95-0, 145-2, 183-0 und 318-2. Die genomische Sequenz der *GapB* von *Mesostigma* wird von drei Introns unterbrochen (TP2, 3-0 und 24-0), wohingegen die *GapB* von *Chlorokybus* kein Intron besitzt. Umfangreiche Versuche mittels PCR ausgehend von vorhandener genomischer DNA die *GapA* und *GapB* von *Coleochaete scutata* sowie die *GapB* aus *Chara* zu amplifizieren, blieben erfolglos. Alle etablierten Introns der *GapA* und *GapB*-Gene von Grünalgen und Landpflanzen (siehe Abb. 10) werden von den Konsensussequenzen „GT“ und „AG“ (5' bzw. 3' im Intron) flankiert. Ihre Längen variieren zwischen 79 nt (*Chlorokybus GapA* 183-0) und 956 nt (*Chara GapA* 85-2). Selbst bei offensichtlich homologen Introns besteht abgesehen von der Position und Phase des Introns keinerlei Übereinstimmung in Sequenz oder Länge.

Innerhalb der *GapA* der Landpflanzen lassen sich vier konservierte Intronpositionen erkennen, je zwei von ihnen befinden sich im Bereich des Transitpeptids bzw. in der reifen Untereinheit der abgeleiteten Proteinsequenz. Intron 85-2 ist als einzige Position in der Charophyceae *Chara vulgaris* und dem primitivsten Streptophyten *Mesostigma viride* präsent. Dies bedeutet, dass die Charophyceae *Chlorokybus* und die Monokotyledonen *Oryza* und *Zea* dieses Intron sehr wahrscheinlich wieder sekundär verloren haben. Im Bezug auf die drei *Chlorokybus*-spezifischen Introns 116-0, 310-2 und 333-2 kann nicht eindeutig festgestellt werden, ob es sich hierbei um einen nachträglichen Erwerb in *Chlorokybus* oder einen sekundären Verlust dieser Position in allen anderen untersuchten Arten handelt. Da *Chlorokybus* die primitivste Charophyceae ist und die Intronverteilung beim Übergang von Chloro- zu Streptophyta bisher noch nicht untersucht wurde, müßten umfangreiche Analysen der *GapA*-Genstruktur erfolgen, um hierüber eine eindeutige Aussage treffen zu können.

2. Ergebnisse A

	TP1-0	TP2-0	3-0	24-0	40-0	85-2	95-0	111-0	116-0	145-2	160-0	166-1	183-0	246-0	310-2	318-2	333-2
<i>Arabidopsis</i> GapB	+	+	-	-	+	-	+	-	-	+	-	+	+	-	-	+	-
<i>Pisum</i> GapB	+	+	-	-	+	-	+	-	-	+	-	+	+	-	-	+	-
<i>Medicago</i> GapB	+	+	-	-	+	-	+	-	-	+	-	+	+	-	-	+	-
<i>Oryza</i> GapB	+	+	-	-	+	-	+	-	-	+	-	+	+	-	-	+	-
<i>Zea</i> GapB	-	+	-	-	+	-	+	-	-	-	-	+	-	-	-	+	-
<i>Marchantia</i> GapB	+	+	-	-	+	-	+	-	-	+	-	-	+	-	-	+	-
<i>Chlorokybus</i> GapB	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Mesostigma</i> GapB	?	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Arabidopsis</i> GapA1	+	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
<i>Arabidopsis</i> GapA2	+	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
<i>Pisum</i> GapA	+	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
<i>Medicago</i> GapA	+	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
<i>Oryza</i> GapA	+	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
<i>Zea</i> GapA	+	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
<i>Marchantia</i> GapA	?	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
<i>Chara</i> GapA	?	?	?	-	-	+	-	-	-	-	-	-	-	-	-	?	?
<i>Chlorokybus</i> GapA	?	?	?	-	-	-	-	-	+	-	-	-	+	-	+	-	+
<i>Mesostigma</i> GapA	?	?	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
<i>Cladophora</i> GapA	?	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Chlamydomonas</i> GapA	+	+	-	-	-	-	-	+	-	-	+	-	-	+	-	-	-
<i>Volvox</i> GapA	-	+	-	-	-	-	-	+	-	-	+	-	-	+	-	-	-

Abbildung 10: Intronmatrix *GapA* und *GapB*

In dieser Arbeit etablierte genomische Sequenzen sind **fett** hervorgehoben. „TP1“ und „TP2“ sind Intronpositionen, die sich in der Präsequenz befinden und nicht eindeutig homologisierbar sind. Bereiche, über die keine genomische Sequenzinformation vorliegt, sind mit einem „?“ markiert. Homologe Introns zwischen *GapA* und *GapB* sind grau unterlegt. Die Nummerierung der Intronpositionen basiert auf der *gap1* von *Geobacillus stearothermophilus* als Referenz (Acc.-Nr. P00362).

Anders verhält es sich bei Intron 183-0 der *GapA* von *Chlorokybus* und den *GapB* der Landpflanzen. Da ein unabhängiger Erwerb von identischen Introns (gleiche Position und Phase) extrem unwahrscheinlich ist, war dieses Intron höchstwahrscheinlich schon vor der *GapA/B*-Genduplikation vorhanden und ist in *Mesostigma* und den höheren Charophyceen sekundär verloren worden. Ähnliches trifft für die Position 166-1 der Landpflanzen zu, nur dass alle hier untersuchten Charophyceen dieses Intron sowohl in *GapA* als auch *GapB*-Sequenzen sekundär verloren haben.

Die Introns 166-1 und 183-0 sind die einzigen zwischen *GapA* und *GapB* konservierten Positionen der reifen Untereinheit (Abb. 10), was auf eine Präsenz vor der Genduplikation hindeutet. Da die Transitpeptide sowohl in Länge als auch Sequenz selbst unter nah verwandten Landpflanzen kaum konserviert sind, ist ein Vergleich dieser Introns sehr

2. Ergebnisse A

spekulativ. Wahrscheinlich handelt es sich hierbei aber um sehr alte Introns, die bereits vor der Aufspaltung von Chloro- und Streptophyta vorhanden waren, was auch die Präsenz in der Grünalge *Chlamydomonas reinhardtii* erklären würde. Außerhalb des Bereiches des Transitpeptides sind allerdings keine mit *Chlamydomonas* homologisierbaren Intronpositionen zu beobachten.

GapB-Gene enthalten insgesamt acht Introns in der reifen Untereinheit, wovon sechs innerhalb der Landpflanzen konserviert sind. Über den Zeitpunkt der Entstehung von vier Positionen (40-0, 95-0, 145-2 und 318-2) kann allerdings nur gemutmaßt werden, da sich ein sekundärer Verlust in *Chlorokybus* nicht ausschließen läßt. Auffällig sind zwei spezifische Introns am 5' Ende der *GapB*, die nur in *Mesostigma* präsent sind.

2.2.2 GapC/Cp (EC-Nummer 1.2.1.12)

Die zweite für Landpflanzen spezifische GAPDH-Duplikation betrifft die cytosolisch-glykolytische GapC. Ihre duplizierte Form, GapCp, erwarb ein Signalpeptid für den Chloroplasten-Import und ist ein Schlüsselenzym der plastidären Glykolyse in Landpflanzen (Petersen et al. 2003). Auch an diesem Gensystem wurde die Intronverteilung in Grünalgen und Landpflanzen untersucht, wobei die GapCp bisher nur aus Landpflanzen bekannt ist. Neben den aus Datenbanken etablierten Daten wurden mittels PCR genomische Sequenzen von *Marchantia polymorpha*, *Coleochaete scutata* und *Mesostigma viride* amplifiziert (Accession Numbers siehe 9.1). Die homologen Positionen der „reifen Proteinuntereinheit“ sind in Abbildung 11 dargestellt. Da die GapCp die Transitsequenz erst nach der Genduplikation erworben hat, wurden Intronpositionen in diesem Bereich (Bezeichnung „TP1“, „TP2“, „TP3“) in der Intronmatrix nicht berücksichtigt

Das *GapC*-Gen von *Marchantia* besitzt innerhalb der genomischen *GapC*-Sequenz zehn Introns (7-1, 41-0, 77-2, 111-0, 160-0, 180-1, 213-0, 260-2, 288-2 und 318-2), die mit denen der Samenpflanzen absolut konserviert sind. In der hochentwickelten Charophyceae *Coleochaete scutata* konnten sieben Introns nachgewiesen werden (41-0, 77-2, 111-0, 160-0, 213-0, 260-2 und 288-2), die im amplifizierten Bereich bis auf die fehlende Position 180-1 sämtlich mit denen der Landpflanzen identisch sind. Die für die Primererstellung verfügbare cDNA-Sequenz wurde per RT-PCR mit degenerierten GAPDH-Primern etabliert (Petersen et al. 2003). Da das Amplifikat die Introns 7-1 und 318-2 nicht abdeckt, kann über diese Intronpositionen keine Aussage getroffen werden (in Abb. 11 mit „?“ gekennzeichnet). Die

2. Ergebnisse A

Grünalge *Mesostigma viride* verfügt über vier Introns in der *GapC*. Nur zwei von diesen (41-0, 213-0) sind mit Streptophyten homologisierbar, während die anderen beiden für *Mesostigma* spezifisch sind (97-1, 144-0).

Bisher ist es nicht gelungen, gezielt *GapCp*-Sequenzen aus Charophyceen zu etablieren und auch in dem EST-Projekt von *Mesostigma* (Simon et al. 2006) mit mehr als 10.000 Sequenzen sind nur *GapA*, *GapB* und *GapC* enthalten. Der evolutiv älteste Organismus, der eine *GapCp* besitzt, ist das Lebermoos *Marchantia polymorpha*. Die genomische Sequenz enthält neun Introns (TP1, TP2, TP3, 7-1, 41-0, 77-2, 111-0, 160-0 und 180-1). Alle sechs Introns der reifen Untereinheit sind mit denen von *GapCp* und *GapC*-Genen konserviert.

	7-1	41-0	77-2	97-1	105-0	111-0	144-0	160-0	180-1	213-0	217-0	252-0	260-2	288-2	318-2
<i>Arabidopsis GapC1</i>	+	+	+	-	-	+	-	+	+	-	-	-	+	-	-
<i>Arabidopsis GapC2</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	-
<i>Pisum GapC</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Zea GapC4</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Physcomitrella GapC</i>	+	+	+	-	-	+	-	+	+	-	-	-	+	-	+
<i>Marchantia GapC</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Coleochaete GapC</i>	?	+	+	-	-	+	-	+	-	+	-	-	+	+	?
<i>Mesostigma GapC</i>	-	+	-	+	-	-	+	-	-	+	-	-	-	-	-
<i>Arabidopsis GapCp1</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Arabidopsis GapCp2</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Capsicum GapCp</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Pinus GapCp</i>	+	+	+	-	-	+	-	+	+	+	-	-	+	+	+
<i>Physcomitrella GapCp</i>	-	+	+	-	-	-	-	-	+	+	-	-	+	+	+
<i>Marchantia GapCp</i>	+	+	+	-	-	+	-	+	+	-	-	-	-	-	-
<i>Chlamydomonas GapC</i>	+	-	+	-	+	-	-	+	-	-	+	-	-	-	-
<i>Volvox GapC</i>	+	-	+	-	+	-	-	+	-	-	+	+	-	-	-

Abbildung 11: Intronmatrix *GapC* und *GapCp*

In dieser Arbeit etablierte genomische Sequenzen sind fett dargestellt. „?“ markieren Bereiche, über die keine genomischen Informationen vorliegen. Intronpositionen, die bis zur *GapC* von Chlorophyceae zurückreichen, sind grau unterlegt. Die Nummerierung der Intronpositionen basiert auf der *gap1* von *Geobacillus stearothermophilus* als Referenz (Acc.-Nr. P00362).

GapC und *GapCp* verfügen über eine extrem konservierte Genstruktur. In Landpflanzen wird die mit der reifen Untereinheit korrespondierende genomische Sequenz beider Gene von zehn

2. Ergebnisse A

absolut identischen Intronpositionen unterbrochen. Dies belegt, dass sämtliche Introns schon vor der Duplikation vorhanden waren und seitdem nur gelegentlich sekundär verloren gingen, wie z.B. in *Marchantia* die vier 3' gelegenen *GapCp*-Positionen 213-0, 260-2, 288-2 und 318-2. Zwei Intronpositionen (41-0, 213-0) aus Landpflanzen lassen sich bis zur *GapC* von *Mesostigma* zurückverfolgen. Obwohl die Introns 7-1, 77-2 und 160-0 in *Mesostigma* fehlen, belegt die Präsenz in den Chlorophyceae *Chlamydomonas* und *Volvox* einen evolutiv deutlich älteren Ursprung.

3. Diskussion Teil A - Die Evolution der Streptophyta

3.1 Die phylogenetische Stellung von *Mesostigma*

Einzigartige Ereignisse in der Evolution sind exzellente Marker, um die Verwandtschaftsbeziehungen heutiger Organismenlinien zu klären. Analysen von endosymbiontischen Gentransfers ermöglichen die Rekonstruktion sehr früher „Inter-Kingdom“ Verknüpfungen (Harper und Keeling 2003; Petersen et al. 2006a; diese Arbeit siehe Teil-B „Die Evolution komplexer Algen“), wohingegen Genom- oder Genduplikationen einen Einblick in die rezentere „Intra-Kingdom“ Evolution geben (Lynch und Conery 2000; Petersen et al. 2003). In der vorliegenden Arbeit wurde der Zeitpunkt der *GapA/B*-Genduplikation im Bezug auf den Ursprung der Landpflanzen (Streptophyta) untersucht. Die Präsenz einer spezifischen C-terminalen Extension in der GapB erlaubt eine zweifelsfreie Identifizierung des entsprechenden Gens, wobei die verfügbaren Daten eindeutig zeigen, dass die GapB in Samenpflanzen universell vorhanden ist, im Genom von *Chlamydomonas* hingegen fehlt. In dieser Arbeit wurden cDNA und genomische Sequenzen der *GapA* und *GapB* aus dem Lebermoos *Marchantia polymorpha* und der Grünalge *Mesostigma viride* etabliert und analysiert.

GapB-Sequenzen wurden in einer früheren Arbeit bereits in fast allen Gruppen der Charophyceen inklusive der basalen Alge *Chlorokybus* nachgewiesen (Petersen 1999). Damit erfolgte die Duplikation der *GapA* also schon weit vor der Entstehung der Landpflanzen vor etwa 700 Mio. Jahren (Douzery et al. 2004). Ein entscheidender Befund ist die Präsenz der *GapB* in der einzelligen Grünalge *Mesostigma viride*, dessen eindeutige Klassifizierung bis in jüngste Zeit zweifelhaft war (McCourt et al. 2004). Sie wurde früher als Prasinophyt klassifiziert (Mattox und Stewart 1984), jedoch deuten ultrastrukturelle Merkmale eine Gruppierung zusammen mit Charophyceen an (Melkonian 1989). Selbst die bisherigen molekularen Daten waren nicht eindeutig und unterstützen entweder ein ganz basales Abzweigen vor Chlorophyta und Streptophyta (Viridiplantae) auf der Grundlage von Mitochondrien- und Plastiden-Genomen (Lemieux et al. 2000; Turmel et al. 2002) oder eine Gruppierung zusammen mit Charophyceen anhand von Einzelgenanalysen (Bhattacharya et al. 1998; Marin und Melkonian 1999). Die Entdeckung sowie die phylogenetische Stellung der *GapB* in *Mesostigma* liefern eindeutige Belege für die Zugehörigkeit dieser Grünalge zu den Streptophyten (Abb. 9). Die phylogenetischen Analysen legen eine Platzierung an der Basis der Charophyceen nahe und entsprechende AU-Tests lehnen eine Positionierung innerhalb

3. Diskussion A

oder basal zu den Chlorophyta signifikant ab (Daten hier nicht gezeigt; siehe Petersen et al. 2006b). Die vorliegenden Analysen deuten darauf hin, dass die *GapA/B*-Genduplikation nach der Aufspaltung von Chloro- und Streptophyta in einem gemeinsamen Vorfahren von *Mesostigma* und Streptophyta erfolgte. Dies ist kompatibel mit dem Fehlen der *GapB* in den komplett sequenzierten Grünalgen *Chlamydomonas* und *Volvox* (Chlorophyceae; Chlorophyta) und der Tatsache, dass Petersen et al. (2006b) keine *GapB* aus der Ulvophyceae *Cladophora* (Chlorophyta), der Trebouxiophyceae *Chlorella* (Chlorophyta) und der Prasinophyceae *Scherffelia* (Chlorophyta) isolieren konnten. Allerdings wurden kürzlich überraschenderweise in zwei Genomprojekten von nahe verwandten Prasinophyteen der Gattung *Ostreococcus* (*O. tauri* und *O. lucimarinus*) *GapB*-Sequenzen nachgewiesen (Robbens et al.). Bei den Prasinophyceae handelt es sich um Grünalgen der Abteilung Chlorophyta für die eine basale Stellung innerhalb dieser Abteilung angenommen wird (Fawley et al. 2000; Gillou et al. 2004). Ob diese mixotrophen Organismen (Graham und Wilcox 2000) die *GapB* durch lateralen Gentransfer (LGT) erhalten haben oder ob die *GapA/B*-Genduplikation nicht möglicherweise doch in einem gemeinsamen Vorfahren aller Viridiplantae erfolgte, werden in Zukunft verfügbare Genomsequenzen weiterer Grünalgen zeigen.

3.2 Intron/Exon Struktur von GAPDH-Genen

3.2.1 Intronverteilung in GAPDH-Genen der Viridiplantae

Eine interessante Beobachtung der vorliegenden Arbeit ist die hochkonservierte Genstruktur der GAPDH innerhalb der Landpflanzen (Abbildungen 10 und 11). So sind sieben von acht Introns der *GapB* und alle zehn Introns der *GapC* höherer Landpflanzen bereits in *Marchantia polymorpha* vorhanden, wobei dieses Lebermoos auch keine zusätzlichen Intronpositionen besitzt. Da die Lebermoose als die ursprünglichsten Landpflanzen angesehen werden (Graham et al. 2000), wurden die Introns in der GAPDH somit seit 700 bis 800 Millionen Jahren konstant vererbt (Douzery et al. 2004). Die Positionen 166-1 und 183-0, die zwischen *GapA* und *GapB* konserviert sind (Abb. 10), liefern einen guten Beleg dafür, dass beide Introns schon vor der Duplikation der *GapA* vorhanden waren. Das generelle Fehlen der Position 166-1 in Charophyceen deutet auf unabhängige Verluste hin.

Ein generelles Problem ist das vollständige Fehlen von Introns in genomischen PCR-Amplifikaten, wie dies z.B. bei der *GapB* von *Chlorokybus* zu beobachten ist. Sofern im gleichen Labor mit cDNA des Organismus gearbeitet wird, besteht die prinzipielle Gefahr,

3. Diskussion A

dass es sich um eine cDNA-Kontamination handelt und das Gen doch Introns enthält. In diesem Fall kann aber davon ausgegangen werden, dass die genomische Sequenz tatsächlich keine Introns enthält, da frühere GAPDH-Isolierungen aus *Chlorokybus* in einem anderen Labor durchgeführt wurden (persönliche Mitteilung J. Petersen). Weiterhin konnten in parallelen PCR-Experimenten mit der gleichen DNA von *Chlorokybus* vier Intronpositionen in der genomischen *GapA* nachgewiesen werden.

Im vorangegangenen Kapitel (Kap. 3.1) wurde anhand der phylogenetischen Analyse von *GapA* und *GapB*-Genen gezeigt, dass *Mesostigma viride* evolutiv basal zu den Charophyceen gruppiert und damit den primitivsten Streptophyten repräsentiert. Der Vergleich der Intronpositionen in GAPDH-Genen von *Mesostigma*, Charophyceen und Landpflanzen zeigt, dass drei Introns der Landpflanzen (*GapA* 85-2; *GapC* 41-0 und 213-0) bereits in der einzelligen Grünalge existierten und z.B. die Position 85-2 der *GapA* in *Chlorokybus* und den Monokotyledonen (*Zea*, *Oryza*) sekundär verloren gingen (Abb. 10). Die Introns 3-0 und 24-0 der *GapB* sowie 97-1 und 144-0 der *GapC* besitzt keiner der übrigen untersuchten Streptophyten, was auf einen linienspezifischen Erwerb in *Mesostigma* hindeutet.

Bereits in der Chlorophyceae *Chlamydomonas reinhardtii* sind drei der zehn Intronpositionen der *GapC* (7-1, 77-2, 160-0) präsent und somit vermutlich seit über 900 Millionen Jahren (Douzery et al. 2004) konserviert. Insbesondere im 3' Bereich offenbart die Intronmatrix von *GapC* und *GapCp* (Abb. 11) einen sekundären Verlust von Introns in den Genen einiger Streptophyten (*Marchantia GapCp*, *Arabidopsis GapC*). Der Verlust von Introns wird in der Literatur insbesondere durch zwei Modelle erklärt (Long und Cerff 2003; Roy und Gilbert 2006). Ein (i) Verlust einzelner Intronpositionen bei gleichzeitigem vollständigen Erhalt der kodierenden exonischen Sequenz wird auf eine eher seltene exakte genomische Deletion zurückgeführt. Bei (ii) der doppelten Rekombination des genomischen Originals mit einer revers transkribierten Kopie der gespleißten mRNA (Genkonversion) kommt es, da die reverse Transkription im Bezug auf die mRNA am 3' Ende startet, vornehmlich zum Verlust von Introns im 3' Bereich. Dieses Modell beschreibt schlüssig, warum dabei gleich mehrere benachbarte Intronpositionen betroffen sind und könnte die Intronverteilung der *GapC* von *Mesostigma* und der *GapCp* von *Marchantia* erklären.

Eindeutige Beispiele für einen rezenten Erwerb von Introns werden deutlich seltener beobachtet. Positionen, die innerhalb der Streptophyten-Sequenzen zwar relativ konserviert sind, sich aber nicht bis zu Chlorophyta (*Chlamydomonas*, *Volvox*) zurückverfolgen lassen (z.B. *GapA* 85-2, *GapB* 95-0), könnten erst relativ spät in der Evolution durch Intron

3. Diskussion A

Transposition, Transposon Insertion oder Intron Transfer in ein paraloges Gen in genomischen Sequenzen etabliert worden sein (Roy und Gilbert 2006).

Der Vergleich der Intronmatrizes (Abb. 10 und 11) zeigt einen deutlich höheren Grad an Konservierung bei *GapC/Cp* als bei *GapA/B*-Genen. Alle zehn vergleichbaren Intronpositionen der Landpflanzen sind zwischen *GapC* und *GapCp* absolut identisch, während *GapA* und *GapB* neben drei konservierten Positionen jeweils ein (*GapA*) bzw. fünf (*GapB*) genspezifische Introns in den Landpflanzen besitzen. Im Zusammenhang mit dem vermutlichen Fehlen der *GapCp* in Charophyceen (siehe Kapitel 2.2.2) deutet dies darauf hin, dass die *GapA/B* Genduplikation deutlich früher erfolgte als die Duplikation der *GapC*. Letztere könnte somit direkt mit der Entstehung der Landpflanzen korrelieren.

3.2.2 GAPDH-Introns: „Neuerwerb“ vor 900 Millionen Jahren oder Relikte der uralten Genstruktur

Kersanach et al. verglichen 1994 erstmals alle damals verfügbaren Intronpositionen der GAPDH-Gene von Pflanzen, Tieren und Pilzen. Zwölf Jahre später konnte im Rahmen dieser Arbeit die Matrix der GAPDH-Intronpositionen um die komplett sequenzierten Tiere und Pilze erweitert werden, wobei der Schwerpunkt auf einem möglichst umfangreichen Taxonsampling lag. So wurden bei den Tieren Vertreter von den primitiven Seescheiden (*Ciona intestinalis*) bis hin zum Menschen ausgewählt. Aufgrund der vorliegenden Arbeit ist es zudem möglich, Intronpositionen der *GapA*, *GapB* und *GapC* aus der primitiven Landpflanze *Marchantia polymorpha* (Lebermoos), Charophyceen und *Mesostigma viride* sowie der nukleär kodierten plastidären *GapCp* (Meyer-Gauen et al. 1994) der Landpflanzen zu analysieren. Diese wurden zusammen mit weiteren Vertretern der Viridiplantae in die Matrix (siehe Abb. 23 im Anhang) integriert. Da es insbesondere bei Pilzen viele organismenspezifische Intronpositionen gibt, sind einige Arten aus Gründen der Übersichtlichkeit lediglich in der Abbildungsunterschrift aufgeführt. Es wurden in diesem „Intron-Update“ insgesamt 62 verschiedene Intronpositionen der reifen Untereinheit in 71 unterschiedlichen GAPDH-Genen identifiziert (siehe Abbildung 23, Anhang). Hierbei erkennt man Positionen, die jeweils innerhalb der Abteilung der Viridiplantae, Tiere und Pilze konserviert sind, jedoch auch solche, die sich zwischen den Abteilungen (Reichen) homologisieren lassen.

3. Diskussion A

3.2.2.1 Vergleich der Intronpositionen auf Abteilungs-Ebene

Innerhalb der Abteilung der Landpflanzen und Grünalgen (Viridiplantae) konnte eine hohe Konservierung der Intron/Exon Struktur in *GapA* und *GapB* bzw. *GapC* und *GapCp*-Genen beobachtet werden. Zusammen mit Rotalgen und Glaucocystophyta bilden Landpflanzen und Grünalgen das Superensemble „Plantae“ (Cavalier-Smith 1981), deren monophyletischer Ursprung auf Plastiden- und Wirtszellebene eindeutig belegt ist (Rodríguez-Ezpeleta et al. 2005). Der Vergleich der entsprechenden Intronpositionen in Viridiplantae mit den homologen GAPDHs von Rotalgen (*Galdieria*, *Chondrus*, *Cyanidioschyzon*) und dem Glaucocystophyten *Cyanophora paradoxa* (persönliche Mitteilung R. Cerff; Abb. 23 im Anhang) zeigt jedoch, dass kein einziges Intron zwischen den drei Abteilungen konserviert ist. Das gilt sowohl für die in Plantae universell vorhandene plastidäre *GapA* cyanobakteriellen Ursprungs (Petersen et al. 2006b) als auch für die cytosolische *GapC*.

Auffällig ist die Präsenz des Introns 41-0, welches außer in *GapC* und *GapCp*-Genen der Streptophyta (Charophyceae und Landpflanzen) auch in *GapC*-Genen von Tieren und Pilzen zu finden ist (siehe Abb. 23, hellgrauer Rahmen). Hierbei könnte es sich um ein uraltes Intron der cytosolischen *GapC* handeln, welches schon vor der eukaryotischen Radiation (vor ca. 1,1 Mill. Jahren; Douzery et al. 2004) vorhanden war.

3.2.2.2 Vergleich der Intronpositionen von *GapA* und *GapC*-Genen

Im Gegensatz zu den bisherigen orthologen Vergleichen, bei denen die Gene der Speziation entstammen, werden mit *GapA* und *GapC* Introns paraloger Gene verglichen. Die *GapC* der Tiere, Pilze und Landpflanzen stammt wahrscheinlich vom proteobakteriellen Vorfahren der Mitochondrien ab (*gap1*), die *GapA* wurde im Zuge der plastidären Endosymbiose von Cyanobakterien erworben (*gap2*; Martin et al. 1993; Cerff 1995). Beide Gene haben also in Eukaryoten einen unterschiedlichen eubakteriellen Ursprung (Martin et al. 1993). Insgesamt können acht Intronpositionen (3-0, 111-0, 116-0, 145-2, 160-0, 166-1, 310-2 und 318-2) beobachtet werden, die sowohl in *GapA* (+*GapB*) und *GapC* (+*GapCp*) Genen vorhanden sind (Abb. 23, dunkelgrauer Rahmen). Sind diese Intronpositionen also deutlich älter als die Radiation der Eukaryoten?

Die Frage nach dem Alter von Introns wird nach wie vor anhand zweier gegensätzlicher Modelle kontrovers diskutiert. Eine anzestrale Präsenz von Introns noch weit vor der Entstehung des ersten Eukaryoten und den sekundären Verlust in Prokaryoten vertritt die „introns-early“-Hypothese. Im Rahmen dieser Theorie wird intronischen Sequenzen eine entscheidende Bedeutung bei der Entstehung der ersten Gene durch so genanntes „exon

3. Diskussion A

shuffling“ zugeschrieben (Kersanach et al. 1994; Cerff 1995; Long et al. 1996; de Souza et al. 1998; Roy 2003). Hierbei wurden einzelne Domänen aneinandergesetzt und erzeugten funktionelle Proteine. Insbesondere Phase-0 Introns werden unterstützend für diesen modularen Zusammenbau zitiert (de Souza et al. 1998). Nach der „introns-late“-Hypothese entstanden Introns erst spät in der Evolution nach der Aufspaltung in Pro- und Eukaryoten und hatten zum Zeitpunkt ihrer Entstehung keine wichtige Funktion (Cavalier-Smith 1985). Identische Intronpositionen von paralogenen Genen werden im Rahmen dieser Hypothese durch unabhängige parallele Integration in definierte Konsensussequenzen, so genannte „Protosplice-Sites“, der eukaryotischen Genome erklärt (Sverdlov et al. 2004; Roy und Gilbert 2006) und wären somit das Ergebnis zufälliger Konvergenz. Im Gegensatz dazu könnten Intronpositionen, die sowohl in *GapA* als auch in *GapC*-Sequenzen vorkommen, nach der „introns-early“-Hypothese bereits lange vor der Duplikation, die zu *gap1* (*GapC*) und *gap2* (*GapA*) führte, in Prokaryoten präsent gewesen sein und somit uralte Introns darstellen (Kersanach et al. 1994; Cerff et al. 1994).

Insgesamt vier (111-0, 145-2, 160-0 und 318-2) der acht identischen Intronpositionen sind in mehreren *GapA* (+*GapB*) und *GapC* (+*GapCp*) Genen enthalten (Abb. 23, dunkelgrauer Rahmen). Die Positionen 318-2 und 145-2 findet man in *GapC* bzw. *GapCp* sowie *GapB*-Genen, fehlen allerdings in allen *GapA*-Genen von Landpflanzen, Grünalgen, Rotalgen und Glaucocystophyta. Die gemeinsame uralte Präsenz dieser Introns würde also jeweils mindestens vier unabhängiger Intronverluste bedürfen. Alle übrigen identischen Intronpositionen (3-0, 116-0, 166-1 und 310-2) sind nur sporadisch in jeweils einem Vertreter der Viridiplantae oder Tiere (3-0 nur in *Mesostigma*, 116-0 und 310-2 nur in *Chlorokybus*, 166-1 nur in *Caenorhabditis*) enthalten, was bei einem ancestralen Ursprung ebenfalls multiple unabhängige Verluste bedeuten würde.

Nach dem derzeitigen Stand der Wissenschaft kann jedoch nicht ausgeschlossen werden, dass in der langen Evolution der Eukaryoten (ca. 1 Milliarde Jahre) identische Intronpositionen nicht auch mehrfach parallel an „hot-spots“ entstanden, wie dies das „introns-late“ Modell beschreibt. Die in Abbildung 23 erkennbaren acht Positionen von *GapA* und *GapC* (3-0, 111-0, 116-0, 145-2, 160-0, 166-1, 310-2 und 318-2) wären so nur das Ergebnis zufälliger Konvergenz und besäßen keinen gemeinsamen Ursprung.

4. Ergebnisse Teil B - Die Evolution komplexer Algen

Ein zweiter Schwerpunkt dieser Arbeit ist die Isolierung und Charakterisierung von Genen für Enzyme des Primärstoffwechsels aus Haptophyten. Hierbei wurden verschiedene Strategien verfolgt, um an die vollständigen Sequenzen zu gelangen. So konnten zum einen bereits zu Beginn der Arbeit im Rahmen eines EST-Projekts (siehe Kap 4.2.1) einige Sequenzen der Gene für die cytosolische Glykolyse und den plastidären Calvin-Zyklus von *Pavlova lutheri* etabliert werden. In einem parallelen Ansatz wurden für alle untersuchten Gene degenerierte Primer (siehe Kap. 4.2.2) erstellt und mittels PCR aus Haptophyten sowie Rotalgen, Dinoflagellaten und Cryptomonaden als Referenzorganismen homologe Gensonden amplifiziert. Sowohl mit den partiellen EST-Klonen als auch mit spezifischen RT-PCR Amplifikaten wurden vorhandene Genbanken unter homologen und heterologen Bedingungen gescreent.

4.1 *In Silico* Arbeiten

4.1.1 Taxonsampling für phylogenetische Analysen

Da das „Taxonsampling“ entscheidenden Einfluss auf die Aussagekraft phylogenetischer Analysen haben kann, wurden zunächst die Datenbanken (siehe Material und Methoden 7.12.1) intensiv mittels BLAST-Analysen (siehe Kapitel 7.12.2) nach allen verfügbaren Sequenzen der Gene für Glykolyse und Calvin-Zyklus durchsucht. Hierbei standen vor allem die phylogenetisch kaum charakterisierten Rhodophyten und komplexen Algen mit roten Plastiden (Haptophyta, Cryptophyta, Heterokontophyta und Dinophyta) im Mittelpunkt der Untersuchungen. Fehlende Referenzsequenzen wurden in eigenen praktischen Arbeiten mittels Screening aus den komplexen Algen *Pavlova lutheri* und *Prymnesium parvum* (Haptophyta), *Guillardia theta* (Cryptophyta), *Lingulodinium polyedrum* und *Pyrocystis lunula* (Dinophyta) sowie der Rotalge *Chondrus crispus* etabliert. Für Untersuchungen der Genstruktur wurden des Weiteren genomische DNA von *Phaeodactylum tricornutum* (Diatomee; Heterokontophyta), *Bigelowiella natans* (Chlorarachniophyta) und *Mesostigma viride* (Grünalge) verwendet.

Sequenzen von EST- bzw. Genomprojekten der Haptophyten *Isochrysis galbana* und *Emiliania huxleyi*, welche während der zweiten Hälfte dieser Arbeit in Datenbanken verfügbar wurden, konnten neben den eigenen etablierten Sequenzen von *Pavlova lutheri* und

4. Ergebnisse B

Prymnesium parvum für die Analytik verwendet werden. Aufgrund der klar dokumentierten Verwandtschaft von Dinoflagellaten und den nicht photosynthetischen Apicomplexa und Ciliaten, welche das Superensemble Alveolata bilden (Van de Peer und De Wachter 1997; Rodríguez-Ezpeleta et al. 2005), wurden alle verfügbaren Sequenzen dieser Abteilungen gezielt aus den Datenbanken etabliert.

4.1.2 Authentizität

Die Authentizität der in eigenen praktischen Arbeiten etablierten Sequenzen wurde zunächst über BLAST-Analysen (siehe Material und Methoden) überprüft. Gleichzeitig wurden, soweit möglich, der G/C-Gehalt und die „Codon Usage“ (Codon Bias) mit bereits bekannten Sequenzen aus dem jeweiligen Organismus verglichen, um mögliche Kontaminationen ausschließen zu können (Daten nicht gezeigt).

4.1.3 Lokalisierung

Für die aus Datenbanken etablierten und eigene Sequenzen wurde versucht, anhand der abgeleiteten Aminosäuresequenz die Präsenz putativer Transitpeptide und damit die intrazelluläre Lokalisierung zu bestimmen. Dazu wurden die Vorhersageprogramme ChloroP, SignalP, TargetP und PSORT verwendet (siehe Material und Methoden 7.12.3).

4.2 Etablierung von Sequenzen

4.2.1 EST-Projekt von *Pavlova lutheri*

Die Grundlage des EST-Projekts der Arbeitsgruppe Petersen in Kooperation mit Prof. H. Philippe (Université de Montréal) bildete die im Rahmen meiner Diplomarbeit (Teich 2002) erstellte λ ZAPII cDNA-Bank des Haptophyten *Pavlova lutheri*. Unter den insgesamt 1000 vom 5' Ende ansequenzierten Klonen befanden sich unter anderem partielle cDNAs der Enolase (ENO), Triosephosphat Isomerase (TPI) und Pyruvat Kinase (PYK), sowie eine vollständige Sequenz der Fructose-1,6-bisphosphat Aldolase (FBA Klasse I). Abgesehen von der Pyruvat Kinase (PYK), die wir Frau Dr. M.F. Liaud für ihre Forschungsvorhaben zur Verfügung stellten, wurden alle partiellen EST-Klone als homologe Sonden zur Etablierung der vollständigen cDNA-Sequenzen (Screening von Genbanken) verwendet.

4. Ergebnisse B

4.2.2 Degenerierte Primer

In umfangreichen Datenbankrecherchen etablierte Sequenzen der Gene für Glykolyse und Calvin-Zyklus wurden in Alignments zusammengestellt. Dabei konnten in den abgeleiteten Aminosäuresequenzen hoch konservierte Proteinmotive identifiziert werden. In diesen Bereichen sollten Primer erstellt werden, um über RT-PCR Experimente homologe Gensonden zu amplifizieren. Ein Problem stellt hierbei der degenerierte genetische Code dar, der für die meisten Aminosäuren mehrere Codons vorsieht. Dieses wurde durch die Konzipierung so genannter „degenerierter Primer“ (siehe Abb. 12) umgangen, welche ein Gemisch aus verschiedenen Primern, die alle Codon-Möglichkeiten abdecken, sind. Der Grad der Degenerierung hat dabei einen entscheidenden Einfluss auf den Erfolg bei einem PCR-Experiment und sollte den Faktor 1024 nicht überschreiten (persönliche Mitteilung J. Petersen). Da durch frühere Arbeiten (siehe Teich 2002) insbesondere für Haptophyten ein extrem hoher G/C-Gehalt von ca. 65% (>93% in dritter Codon Position; Daten nicht gezeigt) nachgewiesen werden konnte, wurden zur Senkung des Degenerierungsgrades an dritter Position ausschließlich Guanin oder Cytosin für die Primersynthese verwendet.

Ein großer Vorteil degenerierter Primer besteht darin, dass sie nicht auf nur einen Organismus beschränkt sind, sondern theoretisch für ganze Organismengruppen eingesetzt werden können. Zudem besteht die Möglichkeit, Isoformen einzelner Gene in einem Organismus zu amplifizieren.

					100								
AA	T	T	T	V	I	C	L	D	D	Y	H	S	L
nt			ACA	GTA	ATT	TGT	TTA	GAT	GAT	TAT	C		
			C	C	C	C	G	C	C	C			
			G	G	A		C	C					
			T	T			G						
							A						
							T						
P125:	5'	ACS	GTS	ATC	TGC	CTS	GAC	GAY	TAY	C	3'		

Abbildung 12: degenerierte Primer

Konzipierung des degenerierten Primers P125 anhand der konservierten PRK-Sequenz um die Aminosäureposition 100. Nukleotide, die für den Primer verwendet wurden sind rot dargestellt. Bei Verwendung aller möglichen Triplet Varianten ergäbe sich ein Degenerierungsgrad von 6144; unter Verwendung der rot markierten Nukleotide ist er 32-fach degeneriert. AA = Aminosäure, nt = Nukleotid; S = Guanin oder Cytosin, Y = Cytosin oder Thymin.

4. Ergebnisse B

4.2.3 Homologe Gensonden

Für die gezielte Suche in Genbanken werden Gensonden verwendet. Bereits in früheren Arbeiten (Teich 2002) konnte festgestellt werden, dass die Verwendung homologer (identischer) Sonden weitaus einfacher ist als die Verwendung heterologer (ähnlicher) Gensonden. Aus diesem Grund wurde zur Etablierung homologer Sonden eine RT-PCR Strategie mit degenerierten Primern (siehe oben), welche für jedes Enzym der Glykolyse und des Calvin-Zyklus (siehe Abb. 6) konzipiert wurden, verfolgt. Die Amplifikate wurden kloniert und zur Überprüfung ihrer Authentizität (siehe oben) vollständig sequenziert. In Tabelle 1 sind die so gewonnenen Sonden aufgeführt.

Die homologen Sonden wurden für individuelle Screening Ansätze verwendet und im Fall der Phosphoribulokinasen konnten so beispielsweise die Klone aus den Haptophyten *Pavlova* und *Prymnesium*, der Cryptomonade *Guillardia* und der Rotalge *Chondrus* erfolgreich isoliert werden. Allerdings gelang es mit dieser Strategie nicht immer, den gesuchten cDNA bzw. genomischen Klon zu erhalten. Ein prinzipieller Grund hierfür kann eine Kontamination sein, die entweder aus dem verwendeten nicht-akzenischen Algenmaterials stammt oder direkt bei der PCR eingeschleppt wurde. Da aus den Genbanken andere Gene erfolgreich isoliert wurden, können Authentizitätsprobleme der Banken ausgeschlossen werden. Abgesehen von partiellen cDNA-Klonen und heterologen Screeningversuchen, bei denen die Sonden meist schlechter hybridisieren, wurde in der vorliegenden Arbeit beobachtet, dass selbst bei homologen Screenings die Hybridisierungssignale z.T. so schwach sind, dass sie übersehen bzw. von Hintergrundsignalen kaum unterschieden werden können (siehe PRK von *Chondrus crispus*, Kapitel 4.3.4).

4. Ergebnisse B

Gen	Organismus		Anzahl unabhängiger Amplifikate	Länge
6-PGDH	<i>Prymnesium</i>	(H)	1	557bp
G6PDH	<i>Pavlova</i>	(H)	3	281bp
FBP	<i>Prymnesium</i>	(H)	1	581bp
	<i>Lingulodinium</i> *	(D)	1	436bp
ATP-PFK	<i>Prymnesium</i>	(H)	1	318bp
TPI	<i>Pavlova</i> (EST-Klon)	(H)	1	678bp
	<i>Lingulodinium</i>	(D)	1	497bp
PGK	<i>Prymnesium</i>	(H)	1	536bp
	<i>Pavlova</i>	(H)	1	545bp
	<i>Lingulodinium</i>	(D)	1	545bp
	<i>Pyrocystis</i>	(D)	1	554bp
ENO	<i>Pavlova</i> (EST-Klon)	(H)	1	ca. 670bp
PYK	<i>Pavlova</i> (EST-Klon)	(H)	1	ca. 1200bp
TKL	<i>Prymnesium</i>	(H)	1	1004bp
SBP	<i>Prymnesium</i>	(H)	3	380bp
RPE	<i>Prymnesium</i>	(H)	1	416bp
PRK	<i>Prymnesium</i>	(H)	1	350bp
	<i>Pavlova</i>	(H)	1	332bp
	<i>Guillardia</i>	(C)	1	335bp
	<i>Chondrus</i>	(R)	1	335bp

Tabelle 1: Homologe Gensonden

H = Haptophyta, D = Dinophyta, C = Cryptophyta, R = Rhodophyta. 6-PGDH = 6-Phosphoglycerat Dehydrogenase, G6PDH= Glucose-6-phosphat Dehydrogenase, FBP= Fructose-1,6-bisphosphatase, ATP-PFK = ATP-abhängige Phosphofructokinase, FBAI = Fructose-1,6-bisphosphat Aldolase Klasse I, TPI = Triosephosphat Isomerase, PGK = Phosphoglycerat Kinase, ENO = Enolase (2-Phospho-D-Glycerat Hydrolase), PYK = Pyruvatkinase, TKL = Transketolase, SBP = Sedoheptulose-1,7-bisphosphatase, RPE = Ribulose-5-phosphat Epimerase, PRK = Phosphoribulokinase. * Die FBP-Gsonde aus *Lingulodinium* wurde in PCR-Reaktionen mit spezifischen Primern amplifiziert (siehe Kap. 4.3.5.1).

4. Ergebnisse B

4.3 Phylogenetische Untersuchungen von Genen aus komplexen Algen mit Rhodoplasten

Im Anschluss an die umfangreichen Datenbankrecherchen wurden von allen Genen vorläufige phylogenetische Analysen durchgeführt. Anhand dieser Stammbäume wurden die vielversprechendsten Gene zur Aufklärung der evolutiven Beziehung der Wirtszellen oder der Plastiden komplexer Algen ausgewählt. Im weiteren Verlauf der praktischen Arbeit wurde der Fokus auf folgende Gene gelegt: Triosephosphat Isomerase (TPI), Enolase (ENO), Fructose-1,6-bisphosphat Aldolase (FBA), Phosphoribulokinase (PRK), Fructose-1,6-bisphosphatase (FBP) und Sedoheptulose-1,7-bisphosphatase (SBP). Soweit nicht anders erwähnt, handelt es sich bei den im Folgenden dargestellten finalen Analysen um Treefinder Maximum-Likelihood (ML) Analysen (siehe Material und Methoden 7.12.5). Die statistische Unterstützung der Genstammbäume wurde durch Bootstrap (BS) und Edge-Support (ES) Analysen bestimmt.

4.3.1 Triosephosphat Isomerase (TPI)

Triosephosphat Isomerasen (TPI; EC-Nummer 5.3.1.1) katalysieren in Glykolyse und Calvin-Zyklus die reversible Isomerisierung von Glycerinaldehyd-3-phosphat und Dihydroxy-acetonphosphat. Es handelt sich um ein vergleichsweise kleines Protein von nur 240 Aminosäurepositionen, welches in Eukaryoten ein Homodimer bildet.

Von *Pavlova lutheri* wurde im Rahmen des EST-Projekts (Kap. 4.2.1) ein partieller Klon der TPI etabliert. Mittels homologen Screenings konnte die vollständige cDNA aus der λ ZAPII Bank isoliert werden. Die Sequenz umfasst 1115bp (siehe 9.2.1). BLAST-Analysen und Vorhersagen von Lokalisierungsprogrammen (siehe Kap. 4.1.3) deuten an, dass es sich hierbei um das Gen für das cytosolische Enzym handelt.

Durch RT-PCR mit den degenerierten Primern P084 und P085 konnte ein 497 bp großer Bereich der TPI aus *Lingulodinium polyedrum* amplifiziert werden. Da im Laufe der Laborarbeiten jedoch TPI-Sequenzen mehrerer Dinophyta (*Heterocapsa triquetra*, *Amphidinium carterae* und *Alexandrium tamarense*) in Datenbanken veröffentlicht wurden, konnte auf ein Screening in *Lingulodinium* verzichtet werden. Für die phylogenetischen Analysen wurde auf die verfügbaren Daten zurückgegriffen.

4. Ergebnisse B

4.3.1.1 Phylogenetische Analyse der TPI

Abbildung 13 zeigt eine Maximum-Likelihood Analyse mit 73 eukaryotischen und 21 eubakteriellen TPI-Sequenzen. Der Baum enthält alle zurzeit in Datenbanken verfügbaren TPI-Sequenzen von Rotalgen und komplexen Algen (Euglenophyta, Cryptophyta, Dinophyta, Diatomeen und Haptophyta), sowie eine repräsentative Auswahl an Sequenzen von Grünalgen, Landpflanzen, Tieren, Pilzen, Trypanosomen, Oomyceten, Ciliaten und Apicomplexa. Aufgrund ihrer hohen Divergenz wurden archaebakterielle Sequenzen aus der vorliegenden Analyse ausgeschlossen.

Die TPI ist mit nur 240 Aminosäuren ein kleiner Marker, der im Bezug auf die frühen Verwandtschaftsbeziehungen der eukaryotischen Linien keine gute Auflösung besitzt, was aus der phylogenetischen Analyse der TPI deutlich wird (Abb. 13). Der Stammbaum unterteilt sich in zwei Teilstammbäume (Teilstammbaum A und B), wobei sich Teilstammbaum B ausschließlich aus eukaryotischen Sequenzen zusammensetzt und Teilstammbaum A zudem eubakterielle Sequenzen enthält; letzterer fungiert als Außengruppe. Innerhalb des Teilstammbaumes B bilden Grünalgen und Landpflanzen (Viridiplantae), Tiere und Pilze jeweils distinkte Äste. Der pflanzliche Ast enthält cytosolische und plastidäre Sequenzen (nicht gezeigt). Die Grünalge *Chlamydomonas* verfügt nur über eine plastidäre TPI, was auf den in der Literatur bereits beschriebenen Verlust der cytosolischen Glykolyse in Chlorophyceae (Schnarrenberger et al. 1990 und 1994) zurückzuführen ist.

Sowohl die cytosolische als auch die plastidären Sequenzen von *Euglena* gruppieren innerhalb der Rotalgen. Obwohl die Bootstrap-Unterstützung hierfür nur niedrig ausfällt (64%), deutet doch eine gemeinsame Insertion zwischen den Aminosäurepositionen 100 und 101 (nicht gezeigt; Referenzsequenz: cytosolische TPI von *Arabidopsis*, Acc.-Nr. AAA03449) auf einen gemeinsamen Ursprung hin. Da *Euglena* ihren Plastiden durch eine sekundäre Endosymbiose mit einer Grünalge erwarb, wurden die TPIs wahrscheinlich lateral (nicht-endosymbiotisch) akquiriert.

Komplexe Algen mit roten Plastiden bilden mehrere kleinere Gruppierungen (Abb. 13). Inmitten eubakterieller TPIs (Teilstammbaum A) befindet sich ein klar unterstütztes (94-99%) Ensemble von zwei Haptophyten (*Prymnesium* und *Emiliania* 3) und Diatomeen (vier Sequenzen), die von einem gemeinsamen Vorfahren in einem nicht-endosymbiotischen Kontext erworben sein könnten. Eine eindeutige Aussage über eine cytosolische oder plastidäre Lokalisierung dieser Sequenzen ist jedoch anhand der Alignments *in silico* nicht möglich. Im eukaryotischen Teilstammbaum (B) gruppieren komplexe Algen mit Rhodoplasten an drei verschiedenen Positionen. Eine Besonderheit sind hierbei die

4. Ergebnisse B

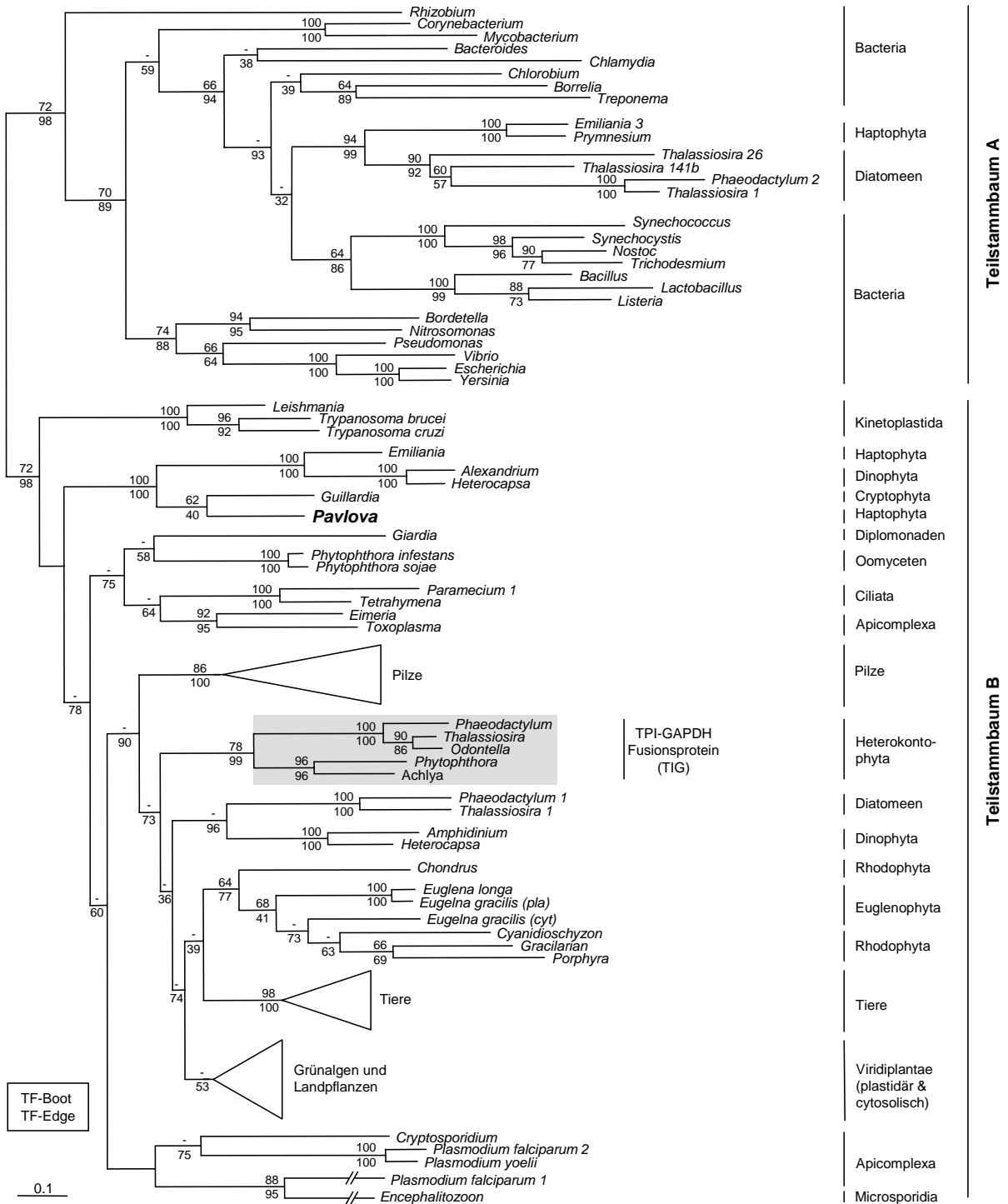


Abbildung 13: Phylogenetische Analyse der Triosephosphat Isomerase (TPI)

Maximum-Likelihood Analyse (Treefinder) mit WAG-Matrix basierend auf 94 TPI-Sequenzen. Die statistische Unterstützung der Äste wurde durch je 100 Bootstrap- (oben) und Edge-Support Analysen (unten) ermittelt. Die im Rahmen dieser Arbeit etablierte TPI aus *Pavlova lutheri* ist **fett** dargestellt.

4. Ergebnisse B

so genannten mitochondrialen TIG-Sequenzen, die einen separaten Ast bilden. Es handelt sich bei der TIG um eine Fusion von TPI und GAPDH, die bisher nur in Heterokontophyta nachgewiesen wurde und Bestandteil einer Glykolyse in den Mitochondrien ist (Liaud et al. 2000; persönliche Mitteilung C. Rio Bartulos). Ein weiteres Ensemble besteht aus cytosolischen TPIs von Dinophyten (*Alexandrium*, *Heterocapsa*), der Cryptomonade *Guillardia* sowie zwei Haptophyten (*Emiliana* und *Pavlova*; 100% BS und ES). *Emiliana* gruppiert hierbei mit 100% (BS und ES) basal zu den beiden Dinophyta während *Pavlova* und *Guillardia* vorher abzweigen, was auf einen lateralen Erwerb der Dinophyten-Sequenzen von Haptophyten hindeuten könnte.

Die nicht-photosynthetischen Apicomplexa findet man an zwei Positionen im eukaryotischen Teilstammbaum (B). Eine dieser Gruppen, bestehend aus *Toxoplasma* und *Eimeria*, läßt eine schwache Affinität zu den Sequenzen der Ciliaten *Paramecium* und *Tetrahymena* erkennen (64% ES), was ihre gemeinsame Zugehörigkeit zu den Alveolata (Van de Peer und De Wachter 1997) widerspiegeln könnte.

4.3.2 Enolase (ENO)

Enolasen (ENO; EC 4.2.1.11) sind essentielle Enzyme der cytosolischen Glykolyse und katalysieren die reversible Mg^{2+} -abhängige Reaktion von 2-Phosphoglycerat zu Phosphoenolpyruvat (PEP). Das PEP plastidärer Enzyme ist eine wichtige Vorstufe für den Shikimat-Weg. Die reife Untereinheit setzt sich aus ca. 440 Aminosäuren zusammen. Analog zur TPI konnte durch das EST-Projekt (siehe Kapitel 4.2.1) ein partieller cDNA-Klon der Enolase von *Pavlova lutheri* etabliert werden. Dieser wurde in einem homologen Screening als Sonde eingesetzt, infolgedessen ein 1594bp großer vollständiger cDNA-Klon isoliert wurde. BLAST-Analysen und Lokalisierungsprogramme (siehe Kap 4.1.3) deuten an, dass es sich hierbei um die cytosolische Isoform des Enzyms handelt.

4.3.2.1 Phylogenetische Analyse der Enolase

Eine Maximum-Likelihood (Treefinder) Analyse der Enolase ist in Abbildung 14 dargestellt. Der Baum enthält insgesamt 92 Sequenzen von Vertretern der Grünalgen, Landpflanzen, Rotalgen, Tiere, Pilze, Ciliaten, Apicomplexa sowie komplexe Algen mit roten (Haptophyta, Heterokontophyta, Cryptophyta, Dinophyta) und grünen Plastiden (*Euglena*, *Bigeloviella*). Die Außengruppe bildet *Euglena* zusammen mit amitochondrialen Protisten und

4. Ergebnisse B

eubakteriellen Sequenzen. Die Gruppierung der rezent duplizierten plastidären und cytosolischen Form von *Euglena* zusammen mit Eubakterien wurde bereits von Hannaert et al. (2000) beschrieben und durch einen horizontalen Gentransfer (HGT) im nicht-endosymbiotischen Kontext erklärt.

Die Analyse der Enolase-Sequenzen (Abb. 14) enthüllt ein sehr komplexes phylogenetisches Muster, da z.B. die Landpflanzen drei distinkte Substammbäume bilden, die aber nicht signifikant unterstützt sind. Lediglich die cytosolischen und plastidären Enolasen (Hannaert et al. 2000) der Landpflanzen gruppieren in einem gemeinsamen Ast, jedoch befinden sich in diesem auch Sequenzen der Oomyceten (z.B. *Phytophthora*) und Haptophyten (*Emiliania* 1, *Prymnesium*, *Isochrysis*). *Bigelowiella*, die ihren Plastiden sekundär von einer Grünalge erwarb, gruppiert in diesem grünen Ast mit geringer statistischer Unterstützung mit den Sequenzen der Haptophyten (BS: <50%; ES: 61%).

Wie bei der phylogenetischen Analyse der TPI beobachtet man auch bei der Enolase an zwei Positionen im Baum eine Gruppierung von Haptophyten und Dinophyten-Sequenzen - eine Beziehung, die offensichtlich einer Verwandtschaft auf 18S rDNA Ebene widerspricht (Dinophyta, Apicomplexa und Ciliata sind Alveolata; Van de Peer und De Wachter 1997). Einen dieser Äste bilden die Sequenzen von *Pavlova*, *Emiliania* 2 (Haptophyta) und *Heterocapsa* 1 (Dinophyta; 100% BS und ES), in einem weiteren Ast gruppieren *Emiliania* 3, *Heterocapsa* 2 und der Dinoflagellat *Karenia* 1, für den bereits ein Erwerb des Plastiden durch tertiäre Endosymbiose eines Haptophyten nachgewiesen ist (Tengs et al. 2000). Letzteres Ensemble gruppiert inmitten cytosolischer, mitochondrialer und plastidärer Sequenzen aus Diatomeen (*Phaeodactylum*, *Thalassiosira*), der Raphidophyceae *Heterosigma* (Heterokontophyta) und den ebenfalls zu den Heterokontophyta zählenden nicht-photosynthetischen Oomyceten *Apodachlya*, *Phytophthora ramorum* und *P. sojae* (90-99%). Aus Rotalgen sind von vier verschiedenen Arten (*Cyanidioschyzon*, *Mastocarpus*, *Prionitis* und *Chondrus*) sechs Sequenzen bekannt. Eine Zuordnung zu bestimmten Kompartimenten (Cytosol, Plastid, Mitochondrium; siehe Kap. 4.1.3) ist jedoch nicht möglich, da es sich bei den verfügbaren Daten um partielle Sequenzen handelt. Basal zu ihnen ordnen sich die zwei Enolase-Sequenzen der Cryptophyten *Guillardia* und *Rhodomonas* an, diese Gruppierung wird jedoch statistisch nicht unterstützt (Edge-Support 35%). Die Sequenzen des Superensembles Alveolata (s.o.) gruppieren basal im Genstammbaum, der Großteil sogar an einem gemeinsamen Ast. Im Falle von *Heterocapsa* 3 und *Karenia* 2 (Dinophyta) handelt es sich um Enolase-GAPDH Fusionstranskripte (Takishita 2004, 2005). Ob diese ein bifunktionelles Enzym bilden oder posttranskriptional/posttranslational getrennt werden, wurde bisher nicht untersucht.

4. Ergebnisse B

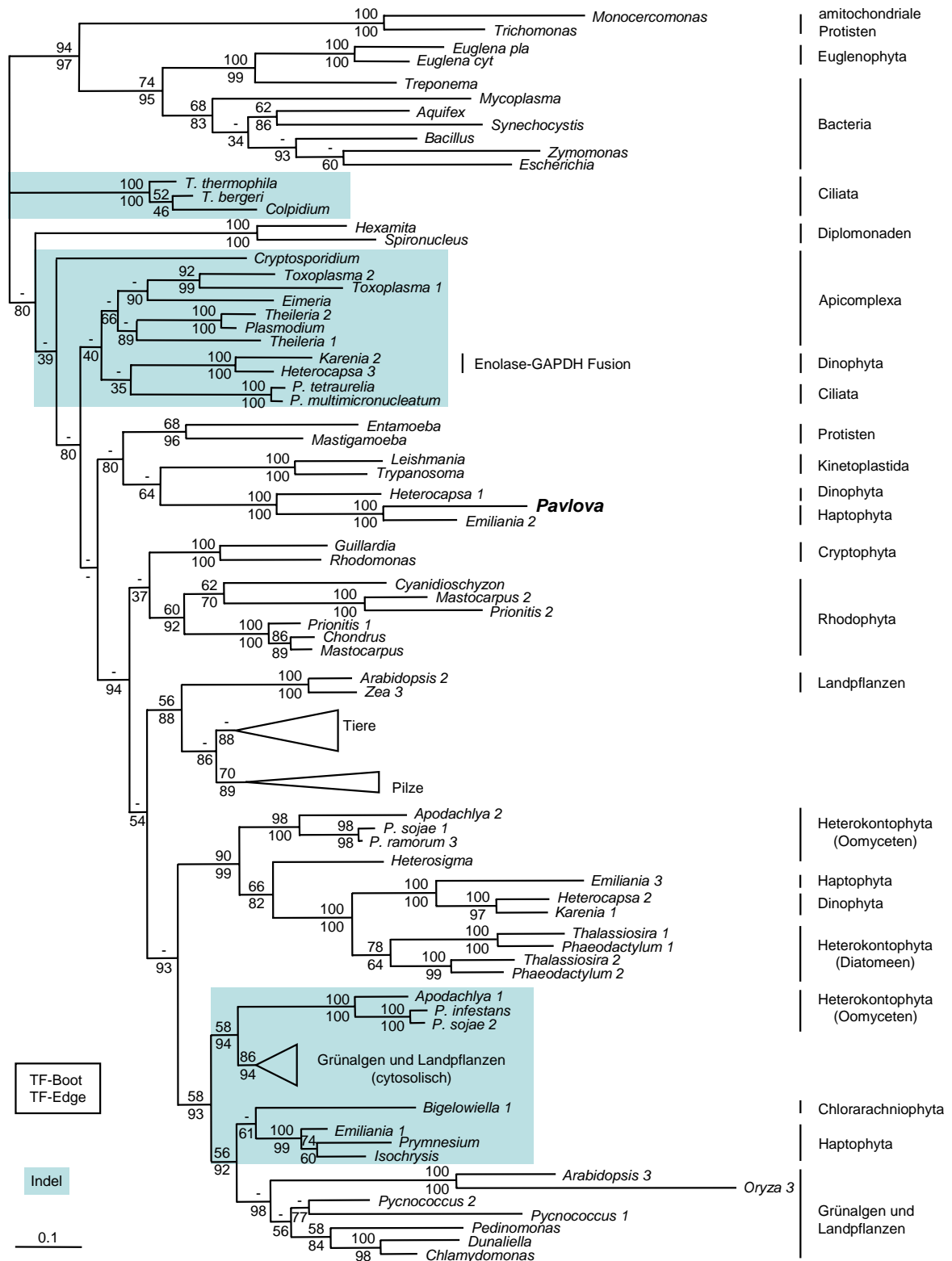


Abbildung 14: Phylogenetische Analyse der Enolase

Maximum-Likelihood Analyse (Treefinder) mit WAG-Matrix basierend auf 92 Enolase-Sequenzen und 412 Aminosäurepositionen. Die statistische Unterstützung der Äste wurde durch je 100 Bootstrap- (TF-Boot; oben) und Edge-Support Analysen (TF-Edge; unten) ermittelt. Die mittels Screening isolierte Enolase aus *Pavlova lutheri* ist **fett** dargestellt. Blau unterlegt sind Sequenzen, die das Indel in den Aminosäurepositionen 103-107 enthalten (siehe auch Abb. 15).

4. Ergebnisse B

Im Alignment fällt insbesondere in einigen Sequenzen von Viridiplantae (Grünalgen und Landpflanzen), Alveolaten, Haptophyten, Oomyceten und *Bigelowiella* eine Insertion von fünf Aminosäuren an den Positionen 103 bis 107 auf (siehe Abb. 15, blau unterlegt; Referenzsequenz für die Nummerierung der Aminosäurepositionen: cytosolische Enolase von *Arabidopsis*, Acc.-Nr. NP181192). Beim Vergleich mit dem ML Baum beobachtet man, dass sich diese Insertion auf zwei distinkte Bereiche beschränkt (Abb. 14, blau unterlegt). Die Verteilung der Insertion korreliert demnach nicht mit der Phylogenie der entsprechenden Sequenzen. Basal im eukaryotischen Ast befindet sich die erste Gruppe bestehend aus Ciliaten, Apicomplexa sowie den beiden Enolasen der Dinophyta *Heterocapsa* 3 und *Karenia* 2 (GAPDH-Enolase Fusionsprotein). Die zweite Gruppe (siehe Abb. 14) umfasst die cytosolischen Sequenzen der Landpflanzen und Grünalgen, drei Oomyceten (*Apodachlya*, zweimal *Phytophthora*), *Bigelowiella* und drei Sequenzen der Haptophyten (*Emiliania*1, *Prymnesium* und *Isochrysis*). Die Proteinsequenz der Insertion ist hoch konserviert und hat in 29 von 34 Fällen das Motiv „XWGWX“ (18/34 „EWGWC“; Abb. 15). Die Aminosäure Tryptophan (W) wird nur von dem Codon „UGG“ kodiert und ist sehr selten. Da in 85% der hier untersuchten Sequenzen diese Aminosäure gleich zweimal verwendet wird, ist ein gemeinsamer evolutiver Ursprung der Indels sehr wahrscheinlich (Dzierszinski et al. 1999; Hannaert et al. 2000 ; Keeling und Palmer 2001; Wilson et al. 2004), Die phylogenetische Analyse (siehe Abb. 14) liefert dafür keine Unterstützung, aufgrund der niedrigen Bootstrap Werte kann diese Schlussfolgerung allerdings auch nicht widerlegt werden.

4. Ergebnisse B

<i>Tetrahymena</i>	KNVTEQEQLDKLMVEQLDGTKN	EWGWCK	SKSLGANAILAVSMAVARAA
<i>Heterocapsa</i> 3	MDVTKQKEIDKLMVETLDGAQN	DWGWSK	ANLGANAILAVSMAVCRAG
<i>Karenia</i> 2	MDVTKQSEIDKLMVETLDGSON	DWGWSK	SNLGANAVLAVSMAICRAG
<i>Paramecium</i>	KNVTEQTKLDKSIVEQLDGSKN	KYGWWS	SKSLGANAILAVSLSLARAG
<i>Plasmodium</i>	MNCTEQKKIDNLMVEELDGSKN	EWGWWS	SKSLGANAILAISMAVCRAG
<i>Heterocapsa</i> 1	KDVCQDEGIDKIML-ELDGTEN	----	SKSLGANAILGVSMVAVCRAG
<i>Pavlova</i>	QDAIDIASLDAQLC-ALDGTDD	----	SKSLGANAILGVSMACRAG
<i>Emiliana</i> 2	MDAQDVAALDKVML-DLDGTPN	----	KGKLGANAILGVSMACRAG
<i>Guillardia</i>	MDPTDQEGIDKKMI-EVDGTPN	----	KTNLGANAILAVSLAVAKAG
<i>Cyanidioschyzon</i>	MDPTEQRAIDERMI-ALDGTPG	GF----	KRRLGANAILGVSMVAVSRAG
<i>Mastocarpus</i> 1	MDASDQGAVDAKMI-ELDGTGG	GF----	KKNLGANAILGVSMVAVSRAG
<i>Chondrus</i>	MDSADQGAIDAKMI-ELDGTGG	GF----	KKNLGANAILGVSMVAVSRAG
<i>Mastocarpus</i> 2	MDPTDQEGIDQALA-DLDGQPD	----	KSRLGANAILGVSMVAVTRAG
<i>Arabidopsis</i> 2	VDVRNQADVDAFML-ELDGTPN	----	SKSLGANAILGVSLSVCRAG
<i>Human</i>	VNVVEQDKIDNLMML-DMDGSEN	----	KSKFGANAILGVSLAVCSNA
<i>Saccarhomyces</i>	IDVKDQKAVDDFLI-SLDGTAN	----	SKSLGANAILGVSLAASRAA
<i>Apodachlya</i> 2	MDPRKQKIDIHML-QLDGTDN	----	KGRLGANAILGVSLAVAKAG
<i>Karenia</i> 1	MSVAEQKALDEKMI-ELDGTPN	----	SKSLGANAILGVSLAATAKAA
<i>Phaeodactylum</i>	MDTADQRAIDQVML-ATDGSPN	----	KANLXANAILGISLAISKAG
<i>Thalassiosira</i> 1	IDVADQRAVDDAML-KADGTPN	----	KGNLGANAILGISLAASKAG
<i>Arabidopsis</i> 1	KDPTQQTAINDFMVHVELDGTQ	N	EWGWCKQKLGANAILAVSLAVCKAG
<i>Oryza</i>	KDPTSQAELDNFMVQQLDGTKN	EWGWCK	QKLGANAILAVSLAICKAG
<i>Zea</i>	KDPTAQTEIDNFMVQQLDGTKN	EWGWCK	QKLGANAILAVSLAVCKAG
<i>Bigelowella</i>	MDPTDQKIDDKMKVELDGSKN	EWGWWS	KSDLGANAILGVSMVAVCRAG
<i>Prymnesium</i>	MDPVQQKEIDDKMKVKTLDGSKN	DWGWWS	SKSLGANAILGVSMVAVCRAG
<i>Isochrysis</i>	MSPIDQKAIDDKMVRELDGSKN	EWGWWS	KAKLGANAILAVSMAVCKAG
<i>Emiliana</i> 1	-----QAIDDKMKVELDGQON	EWGWWS	SKSLGANAILAVSMAVCKAG
<i>Apodachlya</i> 1	KDQNKQTELDTFMVETLDGTKN	EWGWCK	KKKLGANSILAVSLALCKAG
<i>Phytophthora</i>	KDVTQKQAEIDRFMVEKLDGTQ	N	EWGWYKKKLGANAILGVSLALCKAG
<i>Arabidopsis</i> 3	MDPKLQGGIDQAMI-DLDKTEK	----	KSELGANAILAVSIAACKAG
<i>Oryza</i> 3	MDPQQQSQIDQAIM-DLDKAHH	----	KAEIGVNSMLAVSIAACKAG
<i>Chlamydomonas</i>	MDPVKQAEIDQKMK-DLDGTDN	----	KGKLGANAILAVSMAVCKAG

Abbildung 15: Enolase Indel

Ausschnitt aus dem Enolase Alignment. Blau unterlegt ist das konservierte Indel an den Aminosäurepositionen 103-107 (Referenzsequenz: cytosolische Enolase von *Arabidopsis*, Acc.-Nr. NP181192). Die Anordnung der Sequenzen im Alignment entspricht der Anordnung in der phylogenetischen Analyse (Abb. 14).

4.3.3 Fructose-1,6-bisphosphat Aldolase (FBA Klasse I)

Fructose-1,6-bisphosphat Aldolasen repräsentieren einen klassischen Fall von Klasse I/Klasse II Enzymen, die jeweils einen unabhängigen evolutiven Ursprung haben. Sie katalysieren die reversible Aldolkondensation von Glycerinaldehyd-3-phosphat und Dihydroxyacetonphosphat zu Fructose-1,6-bisphosphat im plastidären Calvin-Zyklus und der Gluconeogenese sowie die Spaltung in der Glykolyse und sind somit essentiell für den Primärmetabolismus der Eukaryoten. Im Hinblick auf ihre Reaktionsmechanismen unterscheiden sich beide Formen jedoch deutlich. Klasse I Aldolasen (EC-Nummer 4.1.2.13), die bisher in Pflanzen, Tieren, *Chlamydomonas* und dem Plastiden von *Euglena*

4. Ergebnisse B

nachgewiesen wurden, reagieren mit dem Substrat über die Bildung einer Schiffsschen Base. FBAs der Klasse II benötigen bivalente Kationen als Kofaktor. Klasse II Aldolasen unterteilen sich in Typ A der Ascomyceten und cytosolischen Form von *Euglena* und Typ B von amitochondrialen Protisten, der plastidären FBA aus *Cyanophora paradoxa* und Cyanobakterien (Plaumann et al. 1997; Patron et al. 2004; Kroth et al. 2005). Zu Beginn dieser Arbeit waren aus komplexen Algen mit Rhodoplasten nur plastidäre Enzyme der Klasse II Typ A bekannt. Diese wurden in zwei Diplomarbeiten der Arbeitsgruppe Petersen von Eric Hartleben und Stefanie Verena Junk näher untersucht. Im Rahmen der vorliegenden Arbeit konnte durch das EST-Projekt (siehe Kap. 4.2.1) ein vollständiger cDNA-Klon der cytosolischen FBA Klasse I aus dem Haptophyten *Pavlova lutheri* isoliert werden (1161bp; siehe 9.2.3).

4.3.3.1 Phylogenetische Analyse der FBA Klasse I

Abbildung 16 zeigt eine Maximum-Likelihood Analyse der Klasse I Aldolasen mit 74 eukaryotischen und sieben prokaryotischen Sequenzen. Im Baum sind neben Referenzsequenzen aus Tieren, Viridiplantae (Grünalgen und Landpflanzen) und Rotalgen insbesondere Vertreter der komplexen Algen der roten und grünen Linie (Haptophyta, Diatomeen und Dinophyta; *Euglena* und *Bigelowiella*), Ciliaten, Apicomplexa, Trypanosomen und der Schleimpilz *Dictyostelium* enthalten.

Der Genstammbaum (Abb. 16) spaltet sich in zwei deutlich getrennte Substammbäume, wobei der obere Teilstammbaum Sequenzen der Eubakterien und Diatomeen enthält und sich im unteren überwiegend eukaryotische Sequenzen befinden. Der untere Teilstammbaum beinhaltet einige distinkte und klar unterstützte Äste; einen der Metazoa (97-100%) sowie einen cytosolischen und einen plastidären Ast der Viridiplantae (94-100% Support). In einem weiteren Ast der Viridiplantae befinden sich Sequenzen von Grünalgen und Moosen (BS <50%, ES 56%), deren Lokalisation bisher nicht untersucht wurde. Basal im unteren Teilstammbaum gruppieren Rotalgen mit dem Cyanobakterium *Prochlorococcus* allerdings statistisch nur moderat unterstützt (63-88%). Die Verteilung der Sequenzen komplexer Algen, der Apicomplexa sowie Ciliaten (*Paramecium* und *Tetrahymena*) entspricht keinem offensichtlichen Muster und spiegelt nicht die Verwandtschaftsbeziehung auf 18S rDNA Ebene wider (Van de Peer und De Wachter 1997). Auffällig ist die getrennte Gruppierung der beiden Sequenzen aus Haptophyten; bei ihnen handelt es sich wahrscheinlich um paraloge Gene. *Pavlova* gruppiert mit der cytosolischen FBAI von *Bigelowiella* (64-86%) und *Emiliania* befindet sich in einem Ast der Apicomplexa mit dem Dinophyten *Karenia* (100%).

4. Ergebnisse B

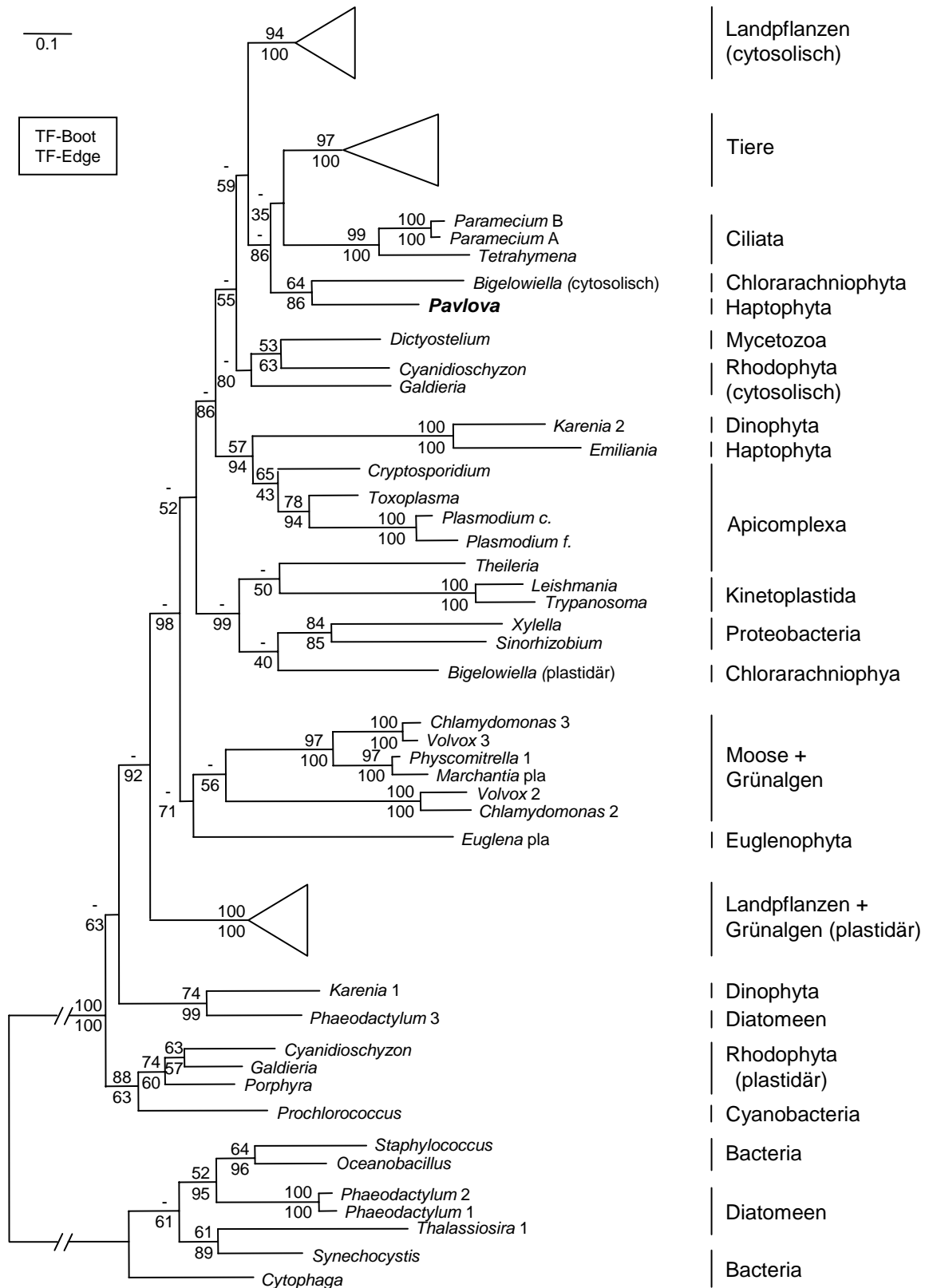


Abbildung 16: Phylogenetische Analyse der FBA Klasse I

Maximum-Likelihood Analyse (Treefinder) mit WAG-Matrix basierend auf 81 FBAI-Sequenzen. Die statistische Unterstützung der Äste wurde durch je 100 Bootstrap- (TF-Boot; oben) und Edge-Support Analysen (TF-Edge; unten) ermittelt. Der vollständige EST-Klon der FBA I aus *Pavlova lutheri* ist **fett** hervorgehoben.

4. Ergebnisse B

Für die Plastiden von *Karenia brevis* wurde bereits ein tertiärer Ursprung von einem Haptophyten nachgewiesen (Tengs et al. 2000), sodass diese Sequenz wahrscheinlich im Zuge einer solchen Endosymbiose von einem Haptophyten aus der Klasse der Prymnesiophyceae (z.B. *Emiliania*; Edvardsen et al. 2000) erworben wurde. Eine weitere Sequenz von *Karenia* befindet sich zusammen mit *Phaeodactylum* (*Phaeodactylum* 3) relativ basal im unteren Teilstammbaum (Support 74-99%).

4.3.4 Phosphoribulokinase (PRK)

Die Phosphoribulokinase (PRK; EC-Nummer 2.7.1.19) regeneriert unter ATP-Verbrauch das CO₂-Akzeptormolekül Ribulose-1,5-bisphosphat und ist eines der Schlüsselenzyme des Calvin-Zyklus. Da es ausschließlich im Plastiden vorkommt, kann nach der Endosymbiose ein funktioneller Ersatz durch eine cytosolische Isoform ausgeschlossen werden. Der Erwerb dieses Gens über den Endosymbionten wäre also das nächstliegende Szenario und die PRK somit ein spezifischer evolutiver Marker für den Plastiden. Ebenso wie bei den Aldolasen existieren zwei Klassen dieses Enzyms. Allerdings findet man PRKs der Klasse I ausschließlich in Proteobakterien, während Cyanobakterien und Eukaryoten das paraloge Enzym der Klasse II besitzen (Martin und Schnarrenberger 1997; Harrison et al. 1998). Anders als die große Untereinheit der Rubisco wurde die PRK bisher wenig untersucht und Sequenzen von komplexen Algen lagen zu Beginn der Arbeit nur aus Heterokontophyta und *Bigeloviella* vor. Durch vorherige Arbeiten in der Arbeitsgruppe Petersen wurden PRK-Sequenzen der Dinoflagellaten *Lingulodinium polyedrum* und *Pyrocystis lunula* sowie von *Euglena gracilis* etabliert. Es war daher das Ziel der vorliegenden Arbeit, Sequenzen aus den beiden verbleibenden Abteilungen komplexer Algen (Haptophyta und Cryptophyta) sowie der Rotalge *Chondrus crispus* zu isolieren. Für RT-PCR Experimente wurden degenerierte Primer konzipiert (siehe Kap. 4.2.2), mittels derer homologe Sonden aus den Haptophyten *Pavlova lutheri* und *Prymnesium parvum*, der Cryptomonade *Guillardia theta* und der Rotalge *Chondrus crispus* amplifiziert werden konnten. In anschließenden homologen Screenings konnten jeweils ein vollständiger cDNA- (*Pavlova* und *Guillardia*) bzw. genomischer Klon (*Prymnesium*) isoliert werden. Anhand der genomischen Sequenz der PRK von *Prymnesium* wurden spezifische Primer erstellt und über RT-PCR Amplifikate die exakte Lage von drei Introns bestimmt. Nachdem ein erster Screeningversuch mit der PRK-Sonde aus *Chondrus*

4. Ergebnisse B

erfolglos war (Thein 2003), konnte in der vorliegenden Arbeit trotz sehr schwacher Hybridisierungssignale ein vollständiger cDNA-Klonen dieser Rotalge identifiziert werden.

4.3.4.1 Charakterisierung der Genstruktur von PRK-Sequenzen

Für den Vergleich der Intronverteilung in Phosphoribulokinasen erfolgten PCR-Experimente mit genomischer DNA. Die Größen von genomischen und cDNA-Amplifikaten wurden zunächst durch Gelelektrophorese verglichen. Bei Größenunterschieden wurden die genomischen Amplifikate eluiert, subkloniert und vollständig sequenziert. Eine exakte Bestimmung der Intronpositionen erfolgte durch den Vergleich von genomischen und cDNA-Sequenzen. Die Intronpositionen des genomischen *Prymnesium* Klons wurden mittels RT-PCR Amplifikaten überprüft. Zur Nummerierung diente die PRK von *Arabidopsis thaliana* (Acc.-Nr. P25697) als Referenz.

Für *Bigelowiella natans*, *Chondrus crispus* und *Guillardia theta* wurden anhand der cDNA-Sequenzen spezifische Primer konzipiert, die genomischen PCR-Produkte kloniert und sequenziert. Das PRK-Gen des Haptophyten *Prymnesium* enthält drei Introns an den Positionen 194-0 (121bp), 244-0 (69bp) und 330-0 (94bp) und der Chlorarachniophyt *Bigelowiella natans* besitzt vier Introns an -49-1 (115bp), 9-0 (67bp), 81-0 (70bp) und 116-0 (109bp). In der genomischen Sequenz der Rotalge *Chondrus crispus* und der Diatomee *Phaeodactylum tricornutum* konnten innerhalb des amplifizierten Bereichs keine Introns nachgewiesen werden. Die etablierte genomische Sequenz der PRK von *Guillardia theta* enthält drei Introns an den Positionen 51-0 (48bp), 116-1 (49bp) und 171-0 (46bp). Beim Vergleich der genomischen PRK-Sequenz von *Guillardia theta* mit dem vorher isolierten cDNA-Klon wurden etliche Sequenzunterschiede insbesondere in dritter Codonposition festgestellt (89% AA-Identität, 92% nt-Identität; siehe 9.2.4). Dies beruht sehr wahrscheinlich auf einer vormals falschen Klassifikation des Organismus, aus dem die verwendete λ NM1149 cDNA-Bank hergestellt wurde. Es handelt sich hierbei höchst wahrscheinlich um *Hanusia phi* (Synonym: *Cryptomonas phi*), einer zu *Guillardia theta* sehr nah verwandten Cryptomonade, was auch in der Diplomarbeit von S.E. Hartleben (2005) anhand der FBA Klasse II beobachtet wurde. Aufgrund dieser nahen Verwandtschaft gelang es mit *Hanusia*-spezifischen Primern erfolgreich den 5' Bereich des PRK-Gens aus *Guillardia* zu amplifizieren. Sämtliche etablierten Intronpositionen sind in einer Intronmatrix zusammengestellt (Abb. 17).

Weiterhin wurden in den Datenbanken verfügbare genomische PRK-Sequenzen analysiert und die Introngrenzen durch den Vergleich mit cDNA-Sequenzen oder den Abgleich mit dem

4. Ergebnisse B

PRK-Alignment bestimmt (siehe Abb. 17). Wie auch *Chondrus crispus* besitzen der primitive Rhodophyt *Cyanidioschyzon merolae* und die Diatomee *Thalassiosira pseudonana* keine Introns in den *PRK*-Genen. Alle Gene der höheren Landpflanzen enthalten vier sowohl in Position als auch Phase (siehe Kap. 1.6) absolut konservierte Introns (175-2, 204-0, 232-1, 314-0), wobei in *Physcomitrella patens* die letzte Intronposition (314-0) nicht vorhanden ist. Die genomischen Sequenzen von *Chlamydomonas reinhardtii* (25-2, 37-0, 60-1, 83-1, 244-0, 299-1, 375-0) und *Volvox carteri* (25-2, 37-0, 60-1, 83-1, 166-0, 244-0, 299-1) werden von sieben Introns unterbrochen. Das Intron 244-0 der beiden Grünalgen findet man auch in den *PRK*-Sequenzen der Haptophyten *Prymnesium parvum* und *Emiliania huxleyi*, wobei dieses Intron in *Emiliania* nicht dem typischen „GT-AG“ Muster (siehe Kap. 1.6) folgt, sondern im 5' Bereich von „GC“ flankiert wird. Darüber hinaus ist dieses Intron auch in dem *PRK*-Gen von *Vaucheria litorea* (Heterokontophyta) vorhanden.

	-49-1	9-0	25-2	37-0	50-0	51-0	60-1	81-0	83-1	116-0	116-1	139-0	161-0	166-0	171-0	175-2	194-0	204-0	232-1	244-0	299-1	314-0	328-0	330-0	375-0
<i>Arabidopsis</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-
<i>Lotus</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-
<i>Oryza</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-
<i>Triticum</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-
<i>Selaginella</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-
<i>Physcomitrella</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-
<i>Volvox</i>	-	-	+	+	-	-	+	-	+	-	-	-	-	+	-	-	-	-	-	+	+	-	-	-	-
<i>Chlamydomonas</i>	-	-	+	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+
<i>Bigelowiella</i>	+	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Emiliania</i>	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	+	+	-	-	-	-	-
<i>Prymnesium</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	+	-
<i>Guillardia</i>	-	-	-	-	-	+	-	-	-	-	+	-	-	-	+	-	-	-	-	-	?	?	?	?	?
<i>Phaeodactylum</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Thalassiosira</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Vaucheria</i>	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	+	-	-
<i>Chondrus</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Cyanidioschyzon</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Abbildung 17: Intronmatrix PRK

Die Nummerierung der Intronpositionen orientiert sich an der *PRK*-Sequenz von *Arabidopsis thaliana* (Acc.-Nr. P25697). **Fett** dargestellt sind genomische Sequenzen, die in dieser Arbeit etabliert wurden. Ein „?“ markiert Positionen, von denen keine genomischen Informationen vorliegen. Die Intronpositionen der *PRK* von *Vaucheria litorea* wurden von Dr. M.E. Rumpho-Kennedy zur Verfügung gestellt.

4. Ergebnisse B

4.3.4.2 Phylogenetische Analyse der PRK

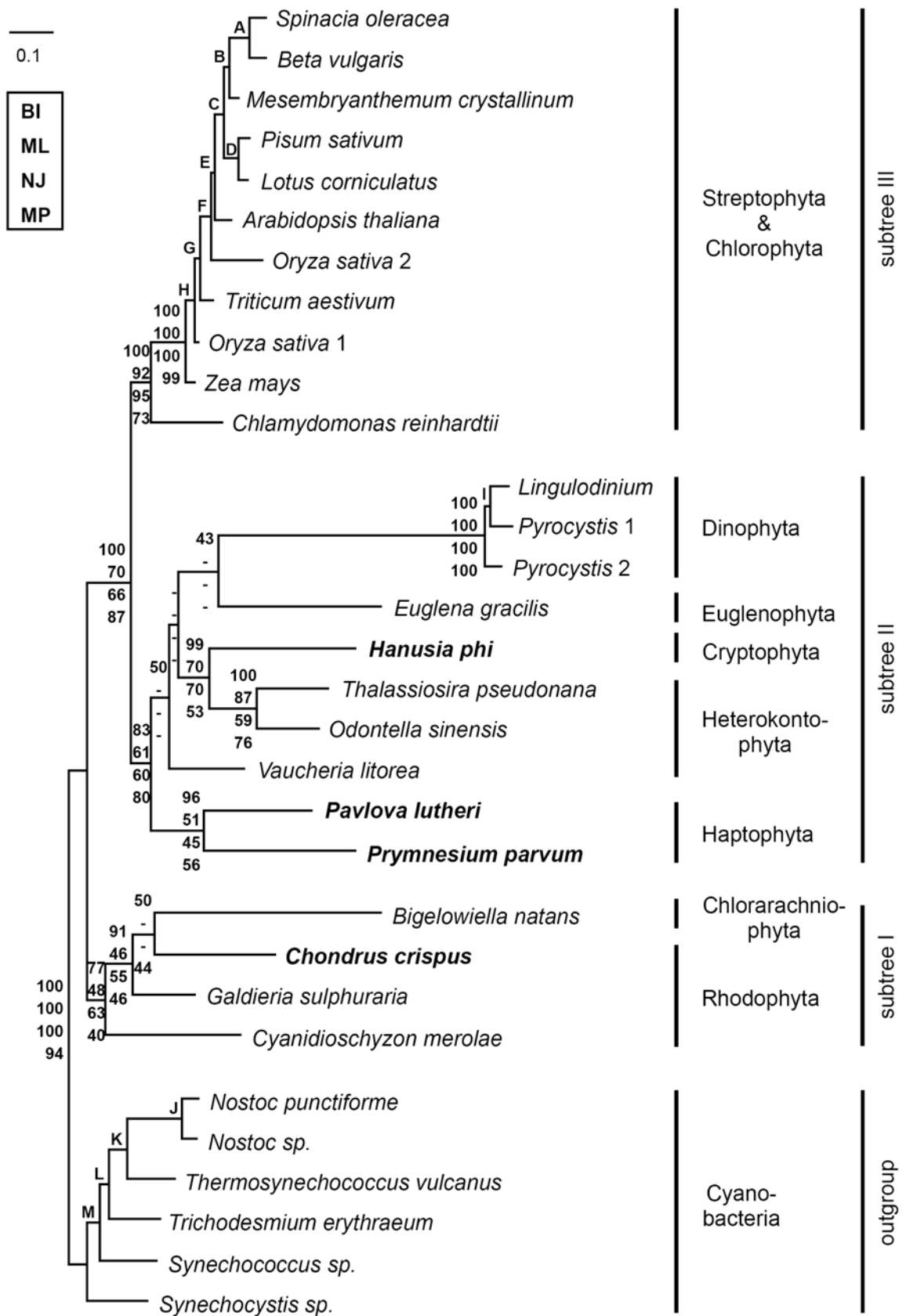
In die phylogenetische Analyse der PRK wurden außer den vier im Rahmen dieser Arbeit etablierten vollständigen Sequenzen aus *Pavlova*, *Prymnesium* (beide Haptophyta), *Hanusia* (Cryptophyta; siehe 4.3.4.1) und *Chondrus* (Rotalge) Daten der Dinoflagellaten *Lingulodinium* und *Pyrocystis* (zwei Sequenzen) und von *Euglena gracilis*, die parallel in der Arbeitsgruppe isoliert wurden, eingeschlossen. Referenzsequenzen stammen aus den Datenbanken.

Abbildung 18 stellt einen MrBayes Consensus Baum mit 25 eukaryotischen und sechs cyanobakteriellen Sequenzen dar, letztere bilden die Außengruppe (outgroup). Er enthält Vertreter der Landpflanzen und *Chlamydomonas*, Rotalgen (*Cyanidioschyzon*, *Galdieria*, *Chondrus*) und Sequenzen aller Algenabteilungen mit komplexen Plastiden (Haptophyta, Dinophyta, Cryptophyta, Heterokontophyta, Euglenophyta und Chlorarachniophyta). Die eukaryotischen Sequenzen bilden drei distinkte Äste. Basal gruppieren die Rotalgen zusammen mit *Bigelowiella* (subtree I) was durch einen hohen Bootstrapsupport unterstützt wird. Die restlichen Sequenzen unterteilen sich in die Algen mit komplexen Rhodoplasten plus *Euglena* (subtree II) und die Landpflanzen und *Chlamydomonas* (subtree III). Innerhalb des Astes II formen die einzelnen Algenabteilungen separate Äste. Lediglich *Vaucheria* (Xanthophyceae; Heterokontophyta) gruppiert basal zu den Diatomeen *Odontella* und *Thalassiosira* (Heterokontophyta), welche zusammen mit *Hanusia* (Cryptophyta) abzweigen. Die Dinophyta besitzen die mit Abstand divergentesten Sequenzen, was sich in einem stark verlängerten Ast widerspiegelt. Nähere Betrachtung des Baumes läßt drei überraschende Gruppierungen erkennen: (i) die Gruppierung der komplexen Algen der roten Linie mit den Landpflanzen und *Chlamydomonas* zum Ausschluss der Rotalgen, (ii) die Affinitäten von *Euglena* zu den komplexen Rotalgen und (iii) *Bigelowiella* innerhalb des Astes der Rotalgen.

Abbildung 18: Phylogenetische Analyse der Phosphoribulokinase (PRK)

MrBayes Consensus Baum (WAG+F+I8+I model) basierend auf 31 PRK-Sequenzen und 292 Aminosäure Positionen. Sequenzen, die in dieser Arbeit etabliert wurden, sind **fett** dargestellt. Support-Werte wurden über Bootstrap-Analysen oder „Posterior Probabilities“ ermittelt. Nur Werte über 30% sind angegeben. BI, Bayesian inference; ML, Maximum-Likelihood; NJ, distance; MP, Maximum-Parsimony. Buchstaben an den Verzweigungen stehen für folgende statistische Support-Werte (BI|ML|NJ|MP): (A) 100|97|98|96; (B) 80|---|---; (C) 82|---|---; (D) 100|90|97|95; (E) 50|---|---; (F) 75|---|---; (G) 62|---|---; (H) 74|---|---; (I) 43|49|65|63; (J) 100|100|100|100; (K) 99|70|60|---; (L) 53|---|---; (M) 78|60|76|---.

4. Ergebnisse B



4. Ergebnisse B

4.3.5 Fructose-1,6-bisphosphatase (FBP) und Sedoheptulose-1,7-bisphosphatase (SBP)

FBP und SBP sind essentielle Enzyme des plastidären Calvin-Zyklus von Landpflanzen (Raines 2003), wo sie über das lichtabhängige Ferredoxin/Thioredoxin-System reguliert werden (Buchanan 1980). Ein cytosolisches Isoenzym der FBP (EC-Nummer 3.1.3.11), welches in Eukaryoten universell präsent ist und aus welchem durch Duplikation das plastidäre Enzym entstand (Martin et al. 1996), katalysiert einen unter physiologischen Bedingungen irreversiblen Schritt der Gluconeogenese (Plaxton 1996). Die komplementäre Reaktion der Glykolyse wird von Phosphofructokinasen katalysiert. Insbesondere die SBP ist für verschiedene Aspekte der Pflanzen- und Evolutionsbiologie bedeutsam. Es wurde gezeigt, dass die SBP den entscheidenden metabolischen „Flaschenhals“ des Calvin-Zyklus darstellt und somit einen bedeutenden Einfluss auf den gesamten Stoffwechselweg besitzt (siehe Review Raines 2003). Dementsprechend erhöht die Überexpression der SBP in Tabak die Nettorate der Kohlenstofffixierung signifikant und bewirkt damit ein schnelleres Pflanzenwachstum, was in der Landwirtschaft in Zukunft verstärkte Bedeutung erlangen könnte (Miyagawa et al. 2001; Lefebvre et al. 2005; Tamoi et al. 2006).

4.3.5.1 Fructose-1,6-bisphosphatase (FBP)

Zu Beginn dieser Arbeit lagen in Datenbanken noch keine Sequenzen der FBP aus Haptophyten vor. Aus isolierter mRNA des charakteristischen Vertreters *Prymnesium parvum* wurde mittels RT-PCR mit degenerierten Primern eine 581bp lange Sequenz der FBP amplifiziert, welche den Bereich von Aminosäure 116 bis 296 umfasst (Referenz aus *Arabidopsis thaliana*: Acc.-Nr. Q9MA79). Da während der praktischen Arbeiten die Sequenz-Rohdaten des Haptophyten *Emiliana huxleyi* freigegeben wurden, konnten zwei (cytosolisch und plastidär) vollständige FBP-Sequenzen durch Datenbankrecherche etablieren werden. Eine phylogenetische Analyse mit dem Amplifikat aus *Prymnesium*, zeigte eine enge Gruppierung mit der plastidären Sequenz von *Emiliana* (nicht gezeigt), sodass auf die Isolierung der vollständigen Sequenzen aus *Prymnesium* verzichtet und stattdessen die verfügbare Sequenz aus *Emiliana* für die phylogenetischen Analysen verwendet wurde (Abb. 19).

Anhand weiterer Datenbankrecherchen konnte eine partielle Sequenz der plastidären FBP aus *Lingulodinium polyedrum* (CD810868) identifiziert werden. Mittels spezifischer Primer (P547/P548; siehe 7.13) und RT-PCR Experimenten gelang es eine homologe Sonde zu

4. Ergebnisse B

amplifizieren, welche für ein erfolgreiches Screening eines nahezu vollständigen cDNA-Klons eingesetzt wurde (ab AA 38; DQ508159). Die plastidäre und cytosolische FBP-Sequenz der Cryptomonade *Guillardia theta* (DQ508162, DQ508158) wurde im Rahmen einer Kooperation von der Arbeitsgruppe Uwe-G. Maier (Philipps-Universität Marburg) zur Verfügung gestellt.

4.3.5.1.1 Phylogenetische Analyse der FBP

Abbildung 19 zeigt eine Maximum-Likelihood Analyse von 94 eukaryotischen und 12 eubakteriellen FBP-Sequenzen. Neben Sequenzen aus Landpflanzen, Grünalgen und Rotalgen sind Vertreter aller Abteilungen komplexer Algen (Cryptomonaden, Haptophyta, Heterokontophyta, Dinophyta, Chlorarachniophyta und Euglenophyta) vorhanden. Weiterhin wurden cytosolische Sequenzen aus Tieren, Pilzen, Kinetoplastida, *Dictyostelium*, *Phytophthora*, *Tetrahymena* und *Toxoplasma* analysiert.

Die Analyse dokumentiert eindeutig den gemeinsamen eubakteriellen Ursprung aller eukaryotischen FBP-Sequenzen. Die Gruppierung von *Toxoplasma* inmitten der Eubakterien deutet auf einen Erwerb durch lateralen Gentransfer (nicht-endosymbiotisch) hin. Innerhalb des Baumes zweigen die plastidären Sequenzen isoliert von den cytosolischen FBPs mit moderater statistischer Unterstützung jedoch einem langen Ast ab (Abb. 19, grau unterlegt). Die Bootstrap-Werte für diese Gruppe steigen auf maximale Werte an (57% zu 97-100%), wenn der Datensatz ohne die schnell evolvierenden plastidären Sequenzen 2 der Diatomeen und Landpflanzen (pla2) analysiert wird (Daten hier nicht gezeigt; siehe Supporting Fig. 1b in Teich et al.). Die übrigen plastidären FBPs (pla1) gliedern sich in einen Ast mit den Viridiplantae (Landpflanzen und Grünalgen) und die Gruppierung der Rhodophyta zusammen mit komplexen Algen (51-65%). Letztere Gruppe enthält sowohl Algen mit komplexen Rhodoplasten (Haptophyta, Cryptophyta, Dinophyta, Heterokontophyta) als auch *Euglena* und *Bigeloviella*, die ihre Plastiden unabhängig voneinander durch sekundäre Endosymbiosen mit Grünalgen erworben haben. Aufgrund der Affinität ihrer plastidären FBPs zu „roten“ Sequenzen ist von einer Rekrutierung dieses Gens in einem nicht-endosymbiotischen Kontext auszugehen (siehe auch PRK, Kapitel 4.3.4).

Der sehr lange Ast der Diatomeen könnte durch ein so genanntes „long branch attraction“ Artefakt (LBA) zu einer falschen Positionierung im Baum führen (Philippe et al. 2000). In der Regel bewirken lange Äste eine Abschwächung der Bootstrap-Unterstützung, was möglicherweise die relativ niedrige Unterstützung für den Ast der komplexen Algen erklärt. Trotzdem entspricht die Assoziation der Diatomeen mit den Rotalgen dem

4. Ergebnisse B

biologischen Erwartungswert und deutet auf einen endosymbiotischen Ursprung der Sequenzen hin.

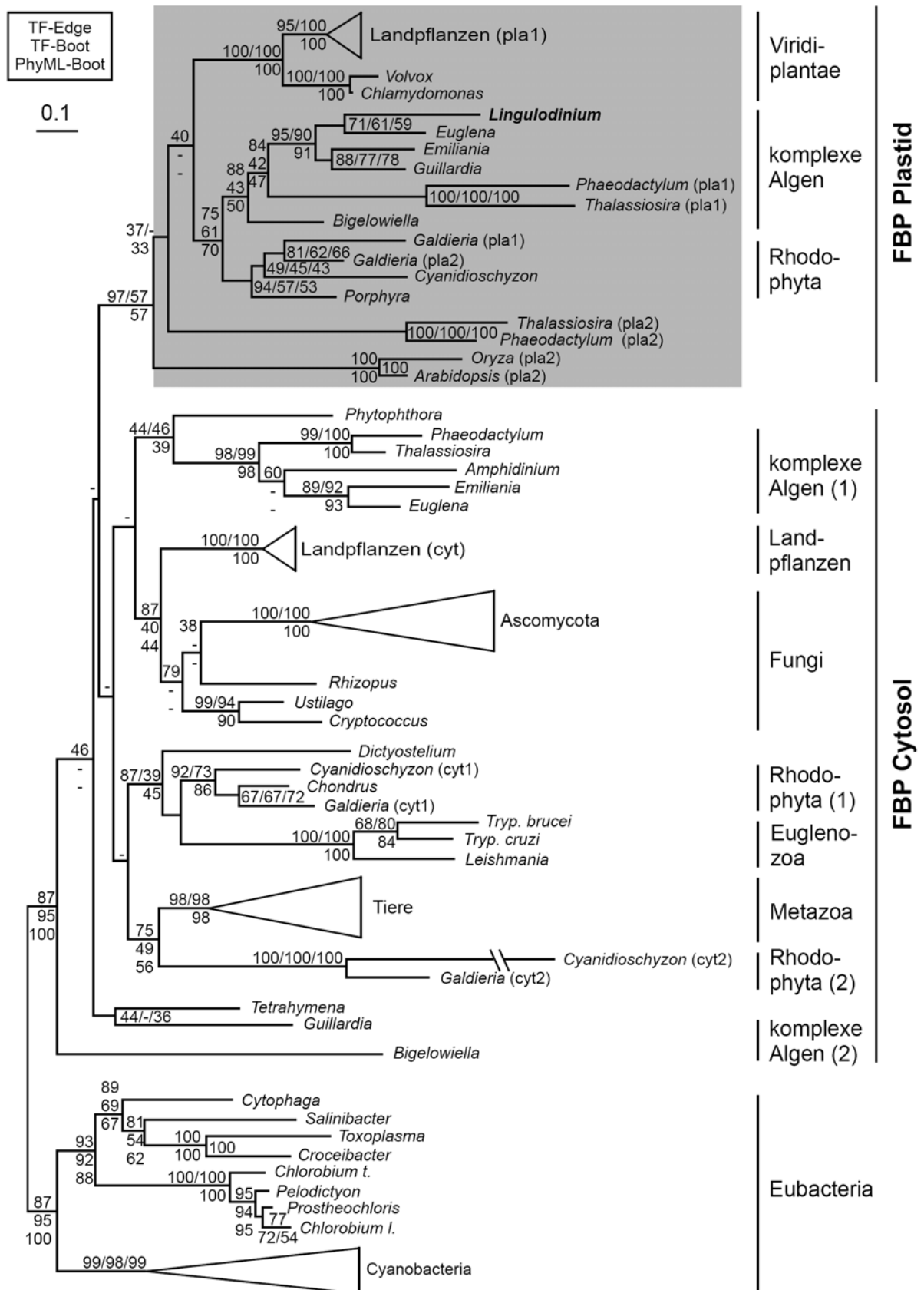
Die cytosolischen FBP-Sequenzen der Landpflanzen, Metazoa und Trypanosomen bilden eindeutige distinkte Äste, wohingegen die Monophylie der Pilze nur schwach unterstützt ist. Durch einen langen Ast sind die Ascomyceten (z.B. *Saccharomyces*) von den Basidio- und Zygomyceten (*Rhizopus*, *Ustilago*, *Cryptococcus*) getrennt. Auch die cytosolischen FBP-Sequenzen der komplexen Algen spiegeln nicht ihre organismische Beziehung wider; *Euglena*, die ihren Plastiden durch sekundäre Endosymbiose mit einer Grünalge erwarb, gruppiert mit den komplexen Algen der roten Linie (98-99% Unterstützung) und nicht mit Trypanosomen. Der gemeinsame Ursprung dieser Sequenzen wird auf Sequenzebene durch zwei spezifische Insertionen von zwei bzw. fünf Aminosäuren (Positionen 289/290 und 384/385; mit *Arabidopsis* CAA41154 als Referenz) unterstrichen. Wie auch bei der plastidären Isoform läßt sich diese Gruppierung mit einem lateralen Erwerb von *Euglena* in einem nicht-endosymbiotischen Kontext plausibel erklären.

In Bezug auf die Evolution der Eukaryoten spiegelt die cytosolische FBP weder die Verwandtschaft von Heterokontophyten (Diatomeen und Oomycet *Phytophthora infestans*) noch das Superensemble Alveolata (Dinophyta, Apicomplexa und Ciliaten) wider. Fehlende statistische Unterstützung der letzten Organismengruppe wird durch die Präsenz offensichtlich paraloger Sequenzen verursacht. Abgesehen von *Toxoplasma* (siehe oben) ersetzte möglicherweise auch *Amphidinium* (Dinophyta) das originale Gen im Zuge der sekundären Endosymbiose.

Abbildung 19: Phylogenetische Analyse der Fructose-1,6-bisphosphatase (FBP)

Maximum-Likelihood Analyse (Treefinder) basierend auf 106 Sequenzen und 231 Aminosäurepositionen. Die in dieser Arbeit etablierte Sequenz von *Lingulodinium* ist **fett** dargestellt. Support-Werte wurden über Bootstrap-Analysen oder „Posterior Probabilities“ ermittelt. Nur Werte über 30% sind angegeben. TF-Edge, Treefinder Edge Support; TF-Boot, Treefinder Bootstrap Werte; PhyML-Boot, PhyML-Bootstrap Werte.

4. Ergebnisse B



4. Ergebnisse B

4.3.5.2 Sedoheptulose-1,7-bisphosphatase (SBP)

Die Sedoheptulose-1,7-bisphosphatase (SBP; EC-Nummer 3.1.3.37) ist ein spezifisches Enzym des plastidären Calvin-Zyklus von Eukaryoten und katalysiert im Rahmen der Regeneration des CO₂-Akzeptors Ribulose-1,5-bisphosphat die Dephosphorylierungsreaktion von Sedoheptulose-1,7-bisphosphat zu Sedoheptulose-7-phosphat. Zu Beginn dieser Arbeit lagen in Datenbanken keine Sequenzen der Haptophyten, Dinoflagellaten und Cryptomonaden vor, weshalb auf diese Sequenzen besonderes Augenmerk gelegt wurde, um für die anschließende phylogenetische Analyse Vertreter aller Algenabteilungen mit komplexen Rhodoplasten zu etablieren.

Durch RT-PCR wurde eine homologe SBP-Sonde des Haptophyten *Prymnesium parvum* amplifiziert, mittels derer im anschließenden Screening ein vollständiger genomischer λ EMBL3-Klon isoliert und sequenziert werden konnte. Die genomische Sequenz beträgt 1851bp und enthält sechs Introns an den Positionen 83-0 (81bp), 130-0 (142bp), 181-1 (68bp), 252-2 (93bp), 300-2 (71bp) & 356-1 (75bp), deren exakte Übergänge ausgehend von spezifischen Primern im Vergleich mit RT-PCR Amplifikaten überprüft wurden. Bis auf Intron 130-0, welches „GC-AG“ Grenzen besitzt, werden alle übrigen Introns von typischen „GT-AG“ Grenzen (siehe Kapitel 1.6) flankiert.

Da es beim Dinoflagellaten *Lingulodinium polyedrum* mittels RT-PCR nicht gelang, eine homologe Gensonde zu etablieren, wurde auf eine heterologe Screeningstrategie mit einem Sondenmix bestehend aus SBP-Klonen von *Prymnesium*, *Euglena*, *Chondrus* und *Porphyra* zurückgegriffen. Zunächst wurde so ein partieller Klon isoliert, mit welchem in einem anschließenden homologen Screening ein vollständiger cDNA-Klon der SBP aus *Lingulodinium* etabliert werden konnte.

Die cDNA-Sequenz der SBP der Cryptomonade *Guillardia theta* (Acc.-Nr. DQ508151) stammt wie auch die FBP-Sequenzen von der Arbeitsgruppe Uwe-G. Maier (Philipps-Universität Marburg).

4.3.5.2.1 Genstruktur der SBP

Analog zur PRK wurde die Intronverteilung der *SBP*-Gene innerhalb photosynthetischer Abteilungen im Bezug auf eine mögliche Korrelation mit der phylogenetischen Analyse (Kap. 4.3.5.2.2) untersucht. Die Intronpositionen der Gene wurden durch den Vergleich von cDNA und genomischer Sequenz exakt bestimmt. Die im Rahmen dieser Arbeit etablierten Sequenzen wurden ausgehend von genomischer DNA und spezifischen Primern mittels PCR

4. Ergebnisse B

gewonnen. Die Nummerierung der Intronpositionen erfolgte im Vergleich mit der einzigen SBP aus *Arabidopsis thaliana* als Referenz (Acc.-Nr. CAB81605).

Zur Überprüfung der Intron Grenzen im genomischen SBP-Klon von *Prymnesium parvum* wurden RT-PCR Amplifikate kloniert und sequenziert. Anhand der cDNA-Sequenzen von *Chondrus crispus* (Rhodophyta; Thein 2003; Acc.-Nr. DQ508156), *Guillardia theta* (Cryptophyta) und *Mesostigma viride* (Chlorophyta) wurden spezifische Primer konzipiert und das Gen mittels PCR-Experimenten mit genomischer DNA amplifiziert. Die Intronpositionen sind in Form einer Intronmatrix zusammengestellt (Abb. 20), wobei in dieser Arbeit etablierte Sequenzen fett hervorgehoben sind. Nicht eindeutig zu homologisierende Positionen innerhalb der nicht konservierten Präsequenz für den Import in den Plastiden wurden in Anführungszeichen gesetzt.

In höheren Landpflanzen (*Arabidopsis* und *Oryza*) findet man sieben Introns („39-0“, 81-0, 143-0, 170-1, 244-1, 278-0, 302-0). Die beiden in Datenbanken gefundenen genomischen SBP-Sequenzen von *Physcomitrella patens* (Laubmoos) enthalten ein Intron („39-0“), sechs der Positionen aus höheren Landpflanzen (bis auf 170-1) sind in dem primitivsten untersuchten Vertreter der Landpflanzen *Marchantia polymorpha* (Lebermoos; Acc.-Nr. DQ873401) vorhanden, was nahe legt, dass beide SBP-Gene von *Physcomitrella* sekundär den Großteil der Introns verloren haben. Da in der einzelligen Grünalge *Mesostigma*, welche zusammen mit Landpflanzen und Charophyceen zu den Streptophyta gehört (siehe Kap. 3.1), das in *Marchantia* fehlende Intron 170-1 nachgewiesen werden konnte, ist die wahrscheinlichste Erklärung, dass dieses Intron in Leber- und Laubmoosen sekundär verloren gegangen ist. Die drei in Landpflanzen im 3' Bereich gelegenen Intronpositionen (244-1, 278-0 und 302-0) fehlen in *Mesostigma*. Eine Konservierung der Intronlängen wurde in keinem Fall beobachtet (nicht dargestellt).

Die beiden nahe verwandten Grünalgen *Chlamydomonas reinhardtii* und *Volvox carteri*, welche zu den ursprünglichen Chlorophyceae (Chlorophyta) gehören, besitzen keines der Introns welches in Streptophyten gefunden wurde. Die SBP-Gene beider Algen enthalten insgesamt sechs Introns, von denen sich fünf an identischen Positionen befinden und jeweils eines spezifisch ist (siehe Abb. 20).

Innerhalb der Algen mit komplexen Rhodoplasten sind keine konservierten Intronpositionen identifizierbar. Während die genomische SBP von *Emiliania huxleyi* kein Intron besitzt, wird die kodierende Sequenz der SBP von *Prymnesium parvum* (beide Haptophyta) von insgesamt sechs Introns (83-0, 130-0, 181-1, 252-2, 300-2 und 356-1) unterbrochen, wobei Intron 130-0 nicht mit der sonst üblichen Sequenz „GT“ sondern „GC“ beginnt (DQ508152).

4. Ergebnisse B

Phaeodactylum tricornutum (Diatomee) enthält nur ein Intron (134-0), *Guillardia theta* (Cryptophyta, DQ508165) besitzt sechs („39-0“, 196-1, 257-0, 296-0, 318-0 und 340-1). Von insgesamt 31 hier identifizierten Intronpositionen wird innerhalb des Bereiches der reifen Proteinuntereinheit nur das Intron 296-0 von zwei Spezies, der Cryptomonade *Guillardia* sowie der *SBP* 2 von *Galdieria* (*Galdieria* 2) geteilt, was eine gemeinsame evolutive Herkunft des Gens widerspiegeln könnte. Das zweite Gen aus *Galdieria* (*Galdieria* 1) besitzt dahingegen ein völlig anderes Intronmuster (vier Introns: „20-0“, „39-0“, 208-1, 271-0). Das im Rahmen dieser Arbeit in *Chondrus crispus* (Rhodophyta; DQ508166) identifizierte Intron (193-2) ist in keiner der beiden *Galdieria*-Sequenzen vorhanden.

	„20-0“	„27-0“	„39-0“	76-0	81-0	83-0	111-0	130-0	134-0	143-0	148-0	170-1	181-1	186-1	193-2	196-1	208-1	212-0	244-1	253-2	256-1	257-0	271-0	278-0	296-0	300-2	302-0	318-0	335-0	340-1	356-1
<i>Arabidopsis</i>	-	-	+	-	+	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-
<i>Oryza</i>	-	-	+	-	+	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-
<i>Selaginella</i>	-	-	-	-	+	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-
<i>Physcomitrella</i> 1+2	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Marchantia</i>	-	-	+	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-
<i>Mesostigma</i>	-	-	+	-	+	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Chlamydomonas</i>	-	+	-	+	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-
<i>Volvox</i>	-	-	-	+	-	-	+	-	-	-	+	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-
<i>Emiliana</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Prymnesium</i>	-	-	-	-	-	+	-	+	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	+
<i>Guillardia</i>	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	+	-	+	-	+	?	-
<i>Phaeodactylum</i>	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Chondrus</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Galdieria</i> 1	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-
<i>Galdieria</i> 2	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-

Abbildung 20: Intronmatrix SBP

Die Nummerierung der Intronpositionen basiert auf der SBP-Sequenz von *Arabidopsis thaliana* als Referenz (CAB81605). **Fett** dargestellt sind genomische Sequenzen, die in dieser Arbeit etabliert wurden. Intronpositionen in Anführungszeichen befinden sich in der Präsequenz und sind nicht eindeutig homologisierbar. Das *SBP*-Gen von *Tetrahymena* (Ciliata) besitzt sechs Introns an den einzigartigen Positionen 113-1, 160-0, 202-2, 241-1, 298-0 und 322-2.

4. Ergebnisse B

4.3.5.2.2 Phylogenetische Analyse SBP

Bereits in früheren Analysen von FBP und SBP-Sequenzen wurde ihr gemeinsamer evolutiver Ursprung dokumentiert und die Monophylie der SBPs nachgewiesen (Martin et al. 1996; Hannaert et al. 2003; Rogers und Keeling 2004). Abbildung 21 zeigt den besten Maximum-Likelihood (ML) Baum (Treefinder), basierend auf der phylogenetischen Analyse von 37 eukaryotischen SBP-Sequenzen. Um LBA-Effekte („long branch attraction“; siehe FBP Kap. 4.3.5.1) zu reduzieren und mehr alignbare Aminosäurepositionen zu erhalten, wurden hoch divergente eubakterielle SBPs und die FBP-Sequenzen aus der Analyse ausgeschlossen. Der Datensatz enthält Sequenzen der primären photosynthetischen Linien (Grünalgen und Landpflanzen, Rhodophyta und *Cyanophora*) und von Vertretern aller komplexen Algen mit roten oder grünen Plastiden. Durch Datenbankanalysen wurden überraschenderweise auch SBPs in den heterotrophen Apicomplexa, Ciliaten, Trypanosomen und Pilzen identifiziert und in die vorliegende Analyse eingeschlossen. Die Wurzel des Baumes bildet die Sequenz von *Cyanophora*.

Die SBP-Phylogenie weist einige distinkte Äste auf, inklusive eines stark unterstützten Astes (96-99%) bestehend aus Viridiplantae und der komplexen Alge *Euglena*, die ihren Plastiden durch Endosymbiose mit einer Grünalge erwarb. Die N- und C-terminalen Kopien von *Euglena* entstammen einer rezenten Genduplikation und werden *in vivo* auf einer cDNA in Reihe und Leseraster (in Frame) konkateniert exprimiert. Mit 86-87%iger Unterstützung zeigen diese eine Affinität zu den Sequenzen der Chlorophyta *Chlamydomonas* und *Volvox*. Im Gegensatz dazu gruppiert die SBP von *Bigelowiella*, welche ebenfalls einen komplexen grünen Plastiden besitzt, zusammen mit den Sequenzen von Rotalgen (72-76% statistische Unterstützung), eine überraschende Beobachtung, die auch bei der PRK-Phylogenie gemacht wurde (siehe Kapitel 4.3.4 und Petersen et al. 2006a) und auf einen lateralen Gentransfer in einem nicht-endosymbiotischen Kontext hindeutet.

Die Sequenzen der Rotalgen sind nicht monophyletisch, was möglicherweise durch unterschiedliche evolutive Raten insbesondere der zusätzlichen SBPs von *Cyanidioschyzon* (B) und *Galdieria* (B) (beide aus der Ordnung Cyanidiales) hervorgerufen werden könnte. An einem langen Ast befindet sich die Gruppierung bestehend aus komplexen Algen mit Rhodoplasten, *Cyanidioschyzon* (C) und vier heterotrophen Linien (Apicomplexa, Ciliaten, Trypanosomen, Ascomyceten). Support Werte von 78-84% liefern nur geringe statistische Unterstützung für einen gemeinsamen Ursprung dieser Sequenzen mit den SBPs der Rotalgen, allerdings könnte die gemeinsame Intronposition 296-0 von *Guillardia* und *Galdieria* (B) darauf hindeuten.

4. Ergebnisse B

Die wahrscheinlich interessanteste Gruppierung der SBP-Phylogenie (Abb. 21) bildet der distinkte Substammbaum mit den Sequenzen aller vier komplexen Algengruppen mit sekundären Rhodoplasten (Diatomeen, Cryptomonaden, Dinophyta und Haptophyta). Basal dazu gruppieren die Sequenzen der beiden heterotrophen Parasiten *Toxoplasma* und *Neospora* (Apicomplexa) mit hoher statistischer Unterstützung (90-95%), welche nach dem Ausschluss der divergenten SBP-Sequenzen von Ascomyceten sogar maximale Werte erreicht (100%; Daten nicht gezeigt). Die SBPs der Apicomplexa könnten so die Gene eines gemeinsamen Rotalgen Endosymbionten repräsentieren, allerdings müßte das Fehlen einer SBP in den komplett sequenzierten Genomen von *Plasmodium*, *Theileria* und *Cryptosporidium* durch einen sekundären Verlust erklärt werden. Innerhalb der Gruppierung der komplexen Algen wird durch moderate Unterstützung (74-79%) eine Verwandtschaft von *Lingulodinium* und den Haptophyten angedeutet, während die basalen Positionen von *Guillardia* und *Phaeodactylum* nicht aufgelöst werden. Die nicht zu diesem Substammbaum gehörende SBP der Diatomee *Thalassiosira* repräsentiert möglicherweise ein paraloges Gen, welches durch eine ursprüngliche Duplikation oder einen lateralen Gentransfer in einem nicht-endosymbiotischen Kontext erworben wurde.

Alle verfügbaren SBP-Sequenzen der nicht photosynthetischen und aplastidären Eukaryoten (Ciliaten, Trypanosomen und Pilze) bilden distinkte und hoch unterstützte Gruppierungen. Der Substammbaum der Pilze (Support 100%) besitzt die längsten Äste, was auf eine deutlich erhöhte Evolutionsrate der entsprechenden Sequenzen hindeutet. Ein Vergleich mit der phylogenetischen Analyse der cytosolischen FBP-Gene offenbart zudem, dass diese SBP-Sequenzen auf eine kleine Untergruppe nahe verwandter Ascomyceten beschränkt sind, wohingegen kein anderes pilzliches Genom SBP-Homologe enthält. Ein lateraler Erwerb der SBP in diesen Ascomyceten ist somit sehr wahrscheinlich. Es fällt zudem auf, dass die SBP-Sequenzen von Diatomeen (*Phaeodactylum* und *Thalassiosira*) ebenso wie die der Kinetoplastida (*Trypanosoma* und *Euglena*) nicht monophyletisch sind. Der unabhängige Ursprung von der SBP aus *Euglena* und *Trypanosomen*, sowie das Fehlen in *Leishmania* (Kinetoplastida, Euglenozoa) deuten auf einen nicht-endosymbiotischen Erwerb der SBP in Trypanosomen hin. Die überraschende Präsenz der SBP in den aplastidären Ciliaten könnte als gemeinsames Merkmal der Ciliaten, Apicomplexa und Dinophyta gewertet werden, jedoch zeigt die phylogenetische Analyse der SBP keine nähere Verwandtschaft des als Alveolata zusammengefaßten Ensembles.

4. Ergebnisse B

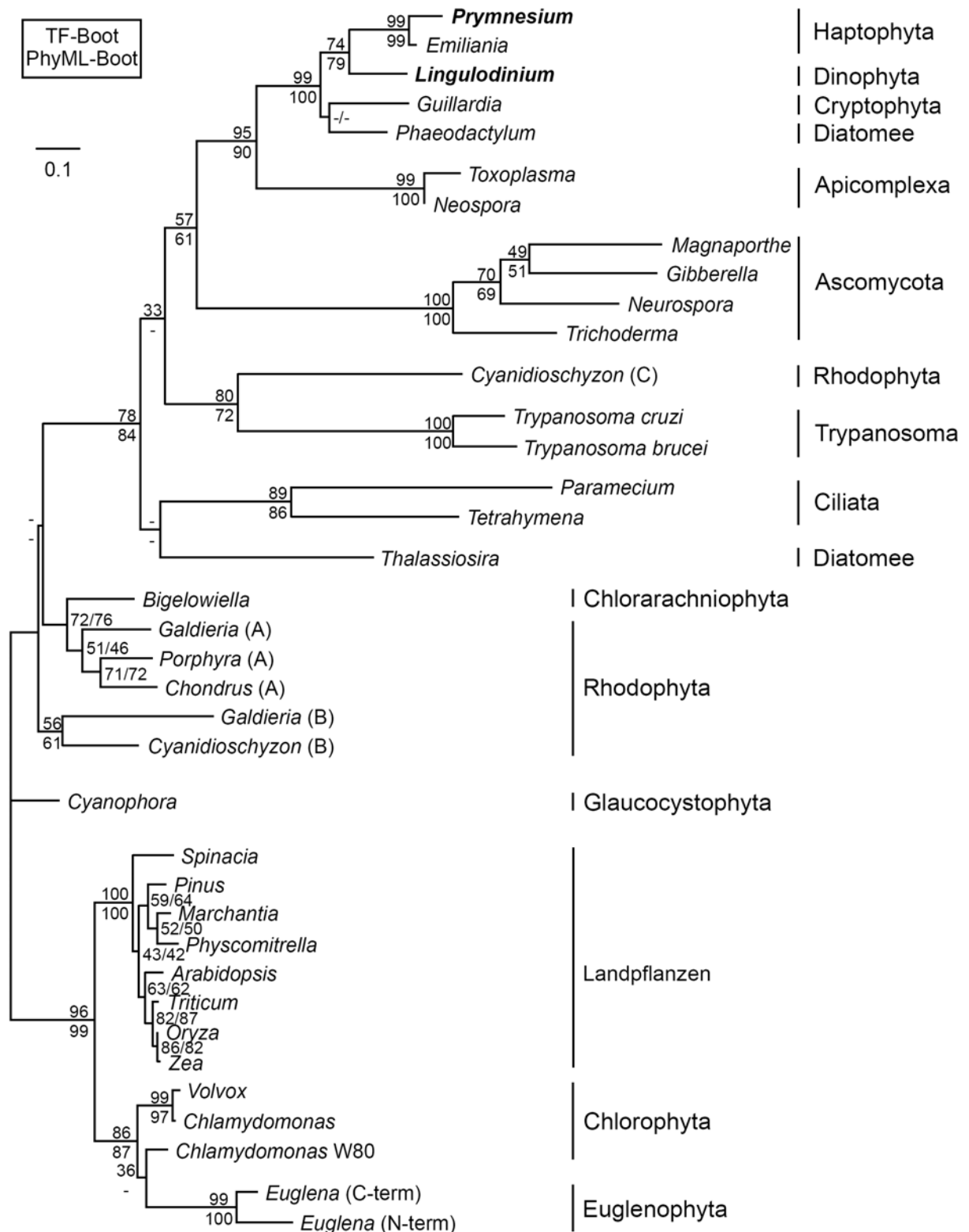


Abbildung 21: Phylogenetische Analyse der Sedoheptulose-1,7-bisphosphatase (SBP)

Treefinder Maximum-Likelihood Baum basierend auf 37 Sequenzen. Sequenzen, die in dieser Arbeit etabliert wurden, sind **fett** dargestellt. Statistische Unterstützung der Topologie erfolgte über Bootstrap-Analysen. Nur Werte über 30% sind angegeben. TF-Boot, Treefinder Bootstrap Werte; PhyML-Boot, PhyML-Bootstrap Werte.

5. Diskussion Teil B - Die Evolution komplexer Algen

„Unser System ist ein Versuch, die Klassifikation mit den neuen, auf molekularbiologische Untersuchungen gegründeten Vorstellungen über die Phylogenese in Einklang zu bringen“

Carl R. Woese

Das Phytoplankton der Weltmeere bildet nur ca. 1% der weltweiten Photosynthese betreibenden Biomasse, ist jedoch für mehr als 45% der jährlichen Nettoprimärproduktion verantwortlich (Field et al. 1998). Die eukaryotischen Algen besitzen dabei einen entscheidenden Einfluss auf die globale Sauerstoffproduktion und den Kohlenstoffkreislauf. Insbesondere durch ihre skurrilen Panzer oder Schuppen aus Calciumcarbonat (Kalk; Dinophyta und Haptophyta) oder Silikat (Diatomeen) faszinieren eukaryotische Mikroalgen, deren fossile Überreste als enorme Kreidablagerungen sichtbar sind und in der Geologie als Leitfossilien dienen. Haptophyten sind sogar an der Wolkenbildung beteiligt (Malin und Kirst 1997) und haben dadurch einen kühlenden Einfluss auf die Erdatmosphäre. Doch auch Massensterben von Muscheln und Fischen können das Ergebnis toxischer Algenblüten von Haptophyten oder Dinophyten sein (Graham und Wilcox 2000).

Trotz ihrer wichtigen ökologischen Stellung ist der evolutive Ursprung von Haptophyten, Heterokontophyten, Cryptomonaden und Dinoflagellaten, die ihre Plastiden durch die sekundäre Endosymbiose mit einer Rotalge erwarben (Douglas 1998; Delwiche 1999), weitgehend unklar. Auf der Basis von 18S rDNA Analysen (Wirtszelle) ist eine Verwandtschaft von Dinoflagellaten mit den nicht-photosynthetischen Apicomplexa und aplastidären Ciliaten gesichert, weshalb diese Gruppe seitdem unter der Bezeichnung Alveolata zusammengefaßt wird (Van de Peer und De Wachter 1997). Die weitreichende Chromalveolaten-Hypothese (CA-Hypothese) postuliert darüber hinaus einen gemeinsamen Ursprung von Alveolaten und den als „Chromista“ bezeichneten Hapto-, Crypto- und Heterokontophyta (Cavalier-Smith 1999). In der vorliegenden Arbeit wurden anhand nukleär kodierter Gene für die cytosolische Glykolyse und den plastidären Calvin-Zyklus die Beziehungen der Wirtszellen und Endosymbionten (Plastiden) von komplexen Algen mit Rhodoplasten untersucht.

5. Diskussion B

5.1 Triosephosphat Isomerase (TPI)

Triosephosphat Isomerasen (TPI) photosynthetischer Eukaryoten sind Bestandteil der cytosolischen Glykolyse und des plastidären Calvin-Zyklus. Die vorliegende Arbeit stellt die bisher umfassendste phylogenetische Analyse der TPI dar (siehe Abb. 13), in der erstmals Vertreter aller Gruppen photosynthetischer Eukaryoten enthalten sind. Ein eubakterieller Ursprung der eukaryotischen TPIs gilt aufgrund der hohen Divergenz archaeabakterieller Sequenzen als gesichert (Keeling und Doolittle 1997; Martin und Schnarrenberger 1997; Ronimus und Morgan 2003). Da in der phylogenetischen Analyse (Abb. 13) weder cytosolische noch plastidäre Sequenzen Affinitäten zu Cyanobakterien aufweisen, legt dies eine proteobakterielle/mitochondriale Abstammung der TPI-Gene nahe (Martin und Schnarrenberger 1997). Die im Zuge der primären plastidären Endosymbiose erworbene cyanobakterielle TPI wurde somit wahrscheinlich durch ein Duplikat des Wirtszell-Enzyms funktionell ersetzt. Eine früher beobachtete nahe Assoziation eukaryotischer TPIs mit α -Proteobakterien (Keeling und Doolittle 1997) wird jedoch von der vorliegenden Analyse (Abb. 13) nicht bestätigt.

Eine Besonderheit in der Analyse stellen die TIGs, eine Fusion aus TPI und GAPDH, dar. Diese konnten bisher ausschließlich in Heterokontophyta nachgewiesen werden (Unkles et al. 1997; Liaud et al. 2000; persönliche Mitteilung C. Rio Bartulos). Aufgrund mitochondrialer Lokalisierung des nukleär kodierten bifunktionellen Enzyms ist es Bestandteil einer möglicherweise ursprünglichen Glykolyse in den Mitochondrien (Liaud et al. 2000).

Die TPI ist mit etwa 240 Aminosäurepositionen ein sehr kleines Protein, weshalb die statistische Unterstützung für die Beziehungen zwischen den eukaryotischen Algenabteilungen in der phylogenetischen Analyse der TPI (Abb. 13) niedrig ist. Sequenzen der komplexen Algen mit Rhodoplasten inklusive der in dieser Arbeit etablierten Sequenz aus *Pavlova lutheri* (Haptophyta) zeigen keinerlei Affinität zu Rhodophyta, sodass das plastidäre Gen nach der sekundären Endosymbiose sehr wahrscheinlich ersetzt wurde. Die cytosolische TPI sollte einen molekularen Marker für die Wirtszelle (der sekundären Endosymbiose) dieser Algengruppe darstellen. Jedoch kann in der vorliegenden Analyse (siehe Abb. 13) weder die Monophylie von Haptophyta, Cryptomonaden und Heterokontophyta (Chromista) noch die der Alveolata (Dinophyta, Apicomplexa, Ciliata) beobachtet werden. Neben des geringen Auflösungsvermögens als phylogenetischer Marker beeinflussen bei der TPI somit möglicherweise auch noch laterale Gentransfers die vorliegende Topologie.

5. Diskussion B

5.2 Enolase (ENO)

Die umfassende phylogenetische Analyse von cytosolischen und plastidären Enolase-Sequenzen (Abb. 14) liefert ein äußerst heterogenes Bild, weshalb Rückschlüsse auf die organismische Evolution äußerst schwierig sind. In früheren Analysen wurden insbesondere so genannte Indels (**I**nsertionen/ **D**eletionen) als evolutive Marker herangezogen (Hannaert et al. 2000; Keeling und Palmer 2001; Baptiste und Philippe 2002; Harper und Keeling 2004). Ein signifikantes Indel von fünf Aminosäuren an den Positionen 103-107 (siehe Abb. 15) ist Bestandteil der Enolase-Sequenzen von Apicomplexa, Ciliaten und Dinophyta, sowie Landpflanzen, Oomyceten, Haptophyta und *Bigelowiella* und wurde in früheren Analysen als Hinweis auf einen „grünen Ursprung“ des heterotrophen Plastiden (Apicoplasten) des Malariaerregers (*Plasmodium*) gedeutet (Dzierszinski et al. 1999), zumal unabhängige parallele Insertionen aufgrund der extremen Sequenz-Konservierung (siehe Abb. 15) und den zwei Tryptophan-Resten (seltene Aminosäure) nahezu auszuschließen sind. Letztlich beschränken sich die Sequenzen mit Indels jedoch auf zwei distinkte Bereiche im Genstammbaum, deren basale bzw. distale Stellung im eukaryotischen Ast (blau unterlegt in Abb. 14) allerdings statistisch nicht unterstützt wird.

Ein Erklärungsmodell (Keeling und Palmer 2001) schlägt einen lateralen nicht-endosymbiotischen Erwerb von einer Grünalge vor, welche dieses Indel in einem Duplikat ihrer ursprünglichen Enolase besaß. Die Autoren prognostizieren aufgrund der distinkten Stellung von Landpflanzen und Alveolata im Stammbaum statt eines vollständigen Replacements eine lokal begrenzte, partielle Rekombination und somit die Entstehung eines Mosaikgens, wie es bereits für Gram-positive Bakterien belegt wurde (Wang und Zhang 2000). Analysen, die sich auf bestimmte Regionen der Sequenz begrenzten und bei denen außerhalb der Region um das Indel keine Affinität von Alveolaten und Landpflanzen festgestellt wurde, unterstützen dieses Modell (Keeling und Palmer 2001). In einer phylogenetischen Analyse des 3' Bereiches konnte jedoch selbst nach der Deletion von 120 Aminosäurepositionen nach dem betreffenden Indel keine signifikante Veränderung der Baumtopologie beobachtet werden (diese Arbeit; Daten nicht gezeigt).

Die vorliegende Arbeit zeigt dieses Indel auch für Haptophyten und den zu den Heterokontophyta zählenden Oomyceten (siehe Abb. 14 und Abb. 15). Zwar würde die reine Präsenz in Haptophyta, Heterokontophyta, Dinophyta (ENO-GAPDH Fusion), Apicomplexa und Ciliaten einen gemeinsamen Ursprung der Wirtszelle unterstützen, jedoch gibt die Topologie des Genstammbaumes dies nicht wieder (siehe Abb. 14; Harper und Keeling 2004),

5. Diskussion B

da insbesondere die Indel enthaltenden Sequenzen der Apicomplexa und Haptophyta weit entfernt voneinander positioniert sind. Das Indel von Haptophyten und Oomyceten müsste aufgrund der vorliegenden phylogentischen Analyse im Rahmen des Modells von Keeling und Palmer (2001) mit zwei weiteren unabhängigen lateralen Gentransfers im nicht-endosymbiotischen Kontext erklärt werden. Die beiden verfügbaren Enolase-Sequenzen aus den Cryptophyten *Guillardia* und *Rhodomonas* besitzen kein Indel.

Da analog zur TPI keine eukaryotische Enolase eine Affinität zu cyanobakteriellen Sequenzen zeigt, könnte in einem alternativen Szenario bereits die Wirtszelle der primären plastidären Endosymbiose ein Duplikat der Enolase mit der Insertion enthalten haben. Nach der Aufspaltung der primär photosynthetischen Linien behielten Viridiplantae (Grünalgen und Landpflanzen) und Rotalgen unabhängig voneinander unterschiedliche Kopien für die cytosolische bzw. plastidäre Funktion. Während die primär grüne Linie weiterhin beide Enolase Kopien besitzt, verloren Rotalgen das Duplikat mit der Insertion. Apicomplexa und Ciliaten hingegen rekrutierten ausschließlich das Enzym mit der Insertion. Dieses Modell erfordert weniger Zusatzannahmen als das von Keeling und Palmer (2001).

Ein aus dem Vergleich von cytosolischen mit plastidären Enolase-Sequenzen resultierendes Paralogieproblem kann bei der Diatomee *Phaeodactylum tricornutum* und dem Haptophyten *Emiliania huxleyi* ausgeschlossen werden, da alle in deren Genomen verfügbaren Sequenzen in die phylogenetische Analyse mit einbezogen wurden (Diatomeen besitzen kein Indel). Wie auch immer dieses Indel letztlich erworben wurde, durch lateralen Gentransfer oder anzestrale Duplikation (was einen paralogen Vergleich bedeutet), kein Szenario setzt einen gemeinsamen Ursprung der Wirtszellgene von Alveolata und „Chromista“ voraus. Da die phylogenetische Analyse dieses Gens keine Monophylie von Haptophyta, Diatomeen und Cryptomonaden („Chromista“) zeigt, liefert dieser Marker kein Indiz für eine evolutive Verwandtschaft dieser Algengruppen.

5. Diskussion B

5.3 Fructose-1,6-bisphosphat Aldolase (FBA)

Die Aldolasen repräsentieren ein sehr komplexes Enzymsystem bestehend aus zwei nicht verwandten Klassen, die darüber hinaus noch in verschiedene Typen unterteilt sind. In früheren Analysen erkannte man zunächst ein vermeintlich einfaches Muster in der Verteilung der unterschiedlichen Enzyme. Eubakterien verfügen über Enzyme beider Klassen (Wittke und Götz 1993). Klasse I Aldolasen fand man zunächst nur als Bestandteil des Stoffwechsels von Tieren, Viridiplantae (Grünalgen und Landpflanzen), Rotalgen und Archaeabakterien, wohingegen Enzyme der Klasse II ausschließlich in Pilzen und Cyanobakterien nachgewiesen wurden (Mutoh und Hayashi 1994; Schnarrenberger et al. 1994; Pelzer-Reith et al. 1995; Martin et al. 1996; Gross et al. 1999). Bereits die Isolierung der Aldolasen aus *Cyanophora paradoxa* und *Euglena gracilis* (Gross et al. 1994; Plaumann et al. 1997) deutete eine komplexere evolutive Geschichte der verschiedenen Enzyme in Eukaryoten an – *Euglena* besitzt ein cytosolisches Enzym der Klasse II und eine plastidäre Aldolase der Klasse I, während die Verteilung in *Cyanophora* reziprok ist. Mittlerweile findet man in den Datenbanken Aldolasen beider Klassen auch in Cyanobakterien, Rotalgen und verschiedenen komplexen Algen mit Rhodoplasten (Kroth et al. 2005). In dieser Arbeit wurden Aldolasen der Klasse I und II analysiert, um einen möglichst vollständigen Datensatz für komplexe Algen zu erhalten.

5.3.1 FBA Klasse II

In zwei Diplomarbeiten von Sven-Eric Hartleben (2005) und Stefanie V. Junk (2005) wurde bereits die Verteilung der Klasse II Aldolasen in komplexen Algen untersucht. Abbildung 22 (siehe Anhang) zeigt einen mit Rotalgen-Sequenzen vervollständigten Maximum-Likelihood Genstammbaum (Treefinder) der FBA Klasse II Typ A. Der Baum lässt sich in einen cytosolischen und einen plastidären Teilstammbaum unterteilen (BS/ES: cytosolisch 71/77%, plastidär 99/99%), wobei letzterer weiter in zwei eindeutig unterstützte Substammbäume (plastidär I und II, 100% Support) verzweigt. Der zweite plastidäre Ast (plastidär II) wird von dem Haptophyten *Emiliania huxleyi* und zwei Diatomeen Sequenzen gebildet, die sehr wahrscheinlich durch eine frühe Duplikation des plastidären Enzyms entstanden sind. Da aus Cryptomonaden und Dinoflagellaten bisher keine komplett sequenzierten Genome verfügbar sind, könnten auch diese Organismen ein homologes Gen besitzen.

Die entscheidende Beobachtung stellt der statistisch eindeutig unterstützte plastidäre Substammbaum I dar (langer Ast; 100% BS und ES), da dieser Sequenzen aus allen

5. Diskussion B

Abteilungen komplexer Algen mit Rhodoplasten enthält und wo mit Ausnahme der Haptophyten alle Organismengruppen distinkte Äste bilden. Es handelt sich hier um einen plastidären Marker, welcher in komplexen Algen mit Rhodoplasten einen gemeinsamen Ursprung hat. Im Gegensatz zu den plastidären Substammbäumen liefern die cytosolischen Sequenzen der Klasse II Aldolasen (Teilstammbaum A) keinen Hinweis auf eine Assoziation von Haptophyten, Heterokontophyten, Cryptomonaden und Dinophyten, was sich mit den Ergebnissen der TPI und Enolase-Analysen deckt.

5.3.2 FBA Klasse I

Bisher waren Klasse I Aldolasen aus Haptophyten gänzlich unbekannt. In der vorliegenden Arbeit wurde aus dem Haptophyten *Pavlova lutheri* (Pavlovophyceae) ein vollständiger cDNA-Klon der FBA Klasse I etabliert und auch in den genomischen TRACE-Files von *Emiliania huxleyi* (Prymnesiophyceae) konnte ein Klasse I Gen nachgewiesen werden, allerdings belegt die phylogenetische Analyse (Abb. 16), dass ihre Gene paralog sind, was auf differentielle Verluste oder einen lateralen Erwerb hindeuten kann. In der phylogenetischen Analyse (Abb. 16) kann kein Ensemble bestehend aus allen Abteilungen komplexer Algen mit Rhodoplasten beobachtet werden, jedoch treten begünstigt durch die Paralogie der Sequenzen aus Haptophyten, sowie die Redundanz von Genen in Diatomeen und *Karenia* (Dinophyta) kleine Gruppierungen dieser Algenabteilungen auf (*Karenia* und *Emiliania*, *Karenia* und *Phaeodactylum* 3). Für *Karenia* wurde bereits ein tertiärer Ursprung des Plastiden durch Endosymbiose eines Haptophyten gezeigt (Tengs et al. 2000), sodass die Gruppierung zusammen mit *Emiliania* auf eine endosymbiotische Rekrutierung aus den Prymnesiophyceae hindeuten könnte. Analog zur cytosolischen Form der Klasse II Enzyme kann in der phylogenetischen Analyse (Abb. 16) kein Ensemble aus Haptophyten, Diatomeen, Dinophyten, Apicomplexa und Ciliaten beobachtet werden.

Obwohl in Datenbanken mittlerweile die Sequenzen zweier Klasse I Enzyme aus den Cyanobakterien *Prochlorococcus* und *Synechocystis* verfügbar sind (siehe Abb. 16), ist ein Erwerb der pflanzlichen Aldolasen Klasse I im Rahmen der plastidären Endosymbiose jedoch unwahrscheinlich, da in Cyanobakterien ansonsten nur Aldolasen der Klasse II nachgewiesen wurden und diese beiden FBA I-Sequenzen keine Affinität zu den pflanzlichen Sequenzen zeigen (Abb. 16). Auch die Verteilung der restlichen Gene korreliert mit keinem Muster, sodass frühe Genduplikationen, laterale Gentransfers und linienspezifische Verluste das gegenwärtige Verteilungsmuster (Abb. 16) erklären können.

5. Diskussion B

5.3.3 Funktionelle Redundanz von Aldolasen?

Die Verteilung von Klasse I Aldolasen (siehe Abb. 16) ist in Viridiplantae und Metazoa (Tiere) von rezenten Duplikationen geprägt, sodass beispielsweise das Genom von *Arabidopsis* sieben verschiedene Gene enthält (nicht dargestellt) und die Diatomee *Phaeodactylum* neben drei Genen der Klasse II weitere drei Gene der Klasse I besitzt (Abb. 16 und 22; Kroth et al. 2005). Es ist unklar, ob es sich bei einigen Aldolase-Duplikaten der Landpflanzen nicht um Pseudogene handelt, oder ob sie einer präzisen Regulation unterliegen, wie dies z.B. für vier unterschiedlich exprimierte cytosolische Glycerinaldehyd-3-phosphat Dehydrogenasen aus Mais, von denen die GapC4 anaerob induziert wird, nachgewiesen wurde (Köhler et al. 1995). Bei Diatomeen findet man generell eine redundante Auslegung des Gen-Sets (nicht gezeigt), wobei bisher nicht bekannt ist, ob es neben der cytosolischen und mitochondrialen Glykolyse sowie dem Calvin-Zyklus noch andere Diatomeen-spezifische Stoffwechselwege gibt, in denen Aldolasen beteiligt sind.

Aus den Rotalgen *Chondrus crispus*, *Gracilaria gracilis* und *Porphyra yezoensis* sind seit kurzem EST-Sequenzen von Aldolasen der Klasse II verfügbar, während diese Gene in den Genomen der primitiven Rotalgen *Cyanidioschyzon merolae* und *Galdieria sulphuraria* (Cyanidiales), die bei pH 0,5-3 und hohen Temperaturen leben (Matsuzaki et al. 2004; Ciniglia et al. 2004), nicht vorkommen, was sich durch einen sekundären Verlust in den basalen Cyanidiales erklären ließe. Insbesondere die durch den Besitz beider Klassen (Rotalgen, Diatomeen) entstandene Redundanz könnte ein Hinweis auf eine feinregulierte Expression sein und eine größere Toleranz gegenüber veränderten Umweltbedingungen ermöglichen, da die Enzyme der beiden Klassen neben ihren unterschiedlichen Reaktionsmechanismen auch verschiedene pH-Optima besitzen. In einer sauren Umgebung wäre die stark pH-Wert abhängige FBaII überflüssig und könnte so z.B. in den Cyanidiales verloren gegangen sein (Kroth 2005).

5.4 Die Phosphoribulokinase (PRK) – ein spezifisches Enzym des Calvin-Zyklus

In der vorliegenden Arbeit wurden PRK-Sequenzen aus Haptophyten (*Pavlova*, *Prymnesium*), Cryptophyta (*Hanusia*) und der Rotalge *Chondrus* etabliert. Zusammen mit weiteren in der Arbeitsgruppe Petersen isolierten cDNAs steht somit erstmals ein Datensatz mit Sequenzen aller Abteilungen photosynthetischer Eukaryoten mit komplexen Plastiden zur Verfügung. Die Phosphoribulokinase (PRK) ist ein spezifisches Enzym des Calvin-Zyklus, was sie zu

5. Diskussion B

einem idealen Testsystem für die Evolution des Plastiden nach Endosymbiosen macht. Für komplexe Algen würde dies im einfachsten Fall bedeuten, dass in einer phylogenetischen Analyse die Gene von Haptophyta, Cryptophyta, Diatomeen und Dinophyta innerhalb der Sequenzen aus Rotalgen und *Euglena*, sowie *Bigelowiella*, die einen komplexen Chloroplasten besitzen (McFadden 1995 und 1997; Durnford 1999; Martin et al 1998; Figge 1999), mit Grünalgen gruppieren.

In der phylogenetischen Analyse der PRK (Abb. 18) fallen die unerwartete Affinität von *Bigelowiella* zu den Rotalgen und die Gruppierung *Euglenas* innerhalb der komplexen Algen mit Rhodoplasten auf. In unserer Publikation der PRK-Daten (Petersen et al. 2006a) wurden die Positionen beider Algensequenzen innerhalb des Baumes mittels „approximately unbiased“ (AU) Tests überprüft. Hierbei verändert man gezielt die Topologie des Baumes („user-defined tree“) und ermittelt anschließend mittels spezieller Programme die Wahrscheinlichkeit dieser Topologie. Sowohl für *Euglena* als auch für *Bigelowiella* wurden jedoch alle möglichen alternativen Platzierungen zusammen mit den Grünalgen signifikant ausgeschlossen (Daten hier nicht gezeigt; siehe Tabelle 1 in Petersen et al. 2006a). Demzufolge haben beide Algen ihre PRK mit großer Wahrscheinlichkeit nach der sekundären Endosymbiose über lateralen, nicht-endosymbiotischen Gentransfer erworben, eine Erklärung, die bereits für eine Reihe anderer Gene von *Bigelowiella* vorgeschlagen wurde (Archibald et al. 2003).

Das Highlight der phylogenetischen PRK-Analyse (siehe Abb. 18) ist die gemeinsame Gruppierung aller komplexen Algen mit Rhodoplasten (66-100% BS). Darüber hinaus überrascht die nahe Assoziation mit den Landpflanzen und der Grünalge *Chlamydomonas*. Die Rotalgen zweigen bereits basal vor diesem Ensemble ab. Diese „grüne Affinität“ komplexer Algen mit Rhodoplast wird zudem in der Genstruktur durch das konservierte Intron 244-0 (siehe Abb. 17), welches in *Chlamydomonas*, *Volvox*, den Haptophyten *Prymnesium* und *Emiliana*, sowie *Vaucheria* (Heterokontophyta; persönliche Mitteilung Dr. M. E. Rumpho-Kennedy) vorkommt, unterstützt. Gleiche Intronpositionen werden generell als Argument für einen gemeinsamen Ursprung der entsprechenden Gene angesehen, weil ein unabhängiger Erwerb an identischen Positionen sehr unwahrscheinlich wäre (Long und Cerff 2003). Die Tatsache, dass das Intron 244-0 innerhalb der primären Linie nur in den Grünalgen *Chlamydomonas* und *Volvox* vorkommt, könnte auf einen nachträglichen Erwerb in Chlorophyta oder im Falle eines ancestralen Introns vor der Aufspaltung in rote und grüne Linie durch parallele Verluste in Landpflanzen und früh in den Rotalgen zurückzuführen sein. Letzteres wird durch das Fehlen des Introns in der basal abzweigenden Rotalge

5. Diskussion B

Cyanidioschyzon angedeutet (siehe Abb. 17). Welches der evolutiven Szenarien letztlich zutrifft, die Präsenz des Introns stimmt mit der Position der Organismen in der phylogenetischen Analyse (Abb. 18) überein. Nach dem Erwerb der „grünen“ PRK in einem gemeinsamen Vorläufer der komplexen Algen mit Rhodoplasten wurde dieses Gen weitervererbt, die Introns jedoch gingen in Cryptophyta und Diatomeen (Heterokontophyta) verloren, wohingegen *Vaucheria* und Haptophyta es noch besitzen.

Wie aber ist die Herkunft der „grünen“ PRK in der komplexen roten Linie zu erklären? Hierfür können mehrere Szenarien in Frage kommen. Genduplikationen (i) sind ein wesentlicher Bestandteil eukaryotischer Evolution. Eine mögliche Duplikation der PRK vor der Aufspaltung der Rhodophyta und Chlorophyta hätte zu zwei Isoformen geführt. Die offensichtlich paraloge Gene entstanden nach der sekundären Endosymbiose durch differentielle Genverluste in den primären Linien. Da bisher in keiner der untersuchten Arten inklusive der bereits komplett sequenzierten Genome (*Thalassiosira*, *Emiliania*) zwei PRKs identifiziert wurden, müssten diese Verluste sehr früh in der Evolution in den Vorläufern der entsprechenden Linien stattgefunden haben. Selbst unter der Annahme eines gemeinsamen Ursprungs der komplexen Algen mit Rhodoplasten würden für dieses Szenario vier unabhängige Ereignisse Voraussetzung sein: eine Genduplikation in einem gemeinsamen Vorfahren von Rot- und Grünalgen, ein Transfer in komplexen Algen mit Rhodoplasten und zwei differentielle Verluste in den primären Linien.

Für ein zweites (ii), weniger umfangreiches Szenario genügt ein einziger nicht-endosymbiotischer lateraler Erwerb im gemeinsamen Vorfahren der komplexen Algen mit Rhodoplasten und der anschließende Ersatz (Replacement) des ursprünglichen Gens der aufgenommenen Rotalge. Solche lateralen Ereignisse sind keine Seltenheit und konnten in dieser Arbeit für *Euglena* und *Bigeloviella* bereits nachgewiesen werden.

Die dritte alternative Erklärung (iii) geht davon aus, dass die Wirtszelle der sekundären Endosymbiose bereits primär photoautotroph war, den grünen Plastiden jedoch wieder verlor (Häuber et al. 1994; Nozaki et al. 2003). In diesem Fall müssten weitere nukleär-kodierte Gene komplexer Algen mit Rhodoplast eine Affinität zu grünen Sequenzen aufweisen, eine Hypothese, über die in der Zukunft der Vergleich ganzer Genome aus komplexen Algen, Rhodophyten, Chlorophyten und Landpflanzen Klarheit geben wird.

Unabhängig davon, welches Szenario die tatsächliche Evolution widerspiegelt, die entscheidende Erkenntnis der phylogenetischen Analyse der PRK ist die Monophylie von Haptophyten, Diatomeen, Cryptophyten und Dinoflagellaten auf der Basis dieses plastidären Markers.

5. Diskussion B

5.5. Fructose-1,6-bisphosphatase (FBP) und Sedoheptulose-1,7-bisphosphatase (SBP)

Nach der 1981 von Weeden formulierten Hypothese („product specificity corollary“) erwartet man nach der plastidären Endosymbiose und einem erfolgreichen Transfer des Gens aus dem Endosymbionten in den Nukleus der Wirtszelle, dass das entsprechende Genprodukt (nach dem Erwerb einer Transitsequenz) wieder in den Plastiden transportiert wird. Erstmals wurde durch die umfangreichen Analysen der Gene des Primärstoffwechsels von Spinat gezeigt, dass neben den erwarteten cyanobakteriellen Proteinen auch Enzyme α -proteobakteriellen Ursprungs in den Plastiden importiert werden (Martin und Schnarrenberger 1997; Martin und Herrmann 1998). Dieser funktionelle Ersatz ist nicht nur auf die Plastiden beschränkt und so wurde z.B. für die cytosolische PGK des Spinats eine cyanobakterielle Herkunft nachgewiesen (Martin und Schnarrenberger 1997). Auch anhand der hier untersuchten FBP und SBP-Sequenzen wird deutlich, dass Weedens Vorhersage häufig nicht zutrifft. Cyanobakterien besitzen neben einer monospezifischen FBP ein bifunktionelles Enzym, welches sowohl Fructose-1,6-bisphosphat als auch Sedoheptulose-1,7-bisphosphat als Substrat akzeptiert (Gerbling et al. 1986). In den Chloroplasten der Landpflanzen wurde diese SBP/FBP mit dualer Substratspezifität durch zwei monospezifische Isoformen eubakteriellen Ursprungs ersetzt. Die nukleär kodierten Gene beider Enzyme sind homolog, besitzen jedoch nur entfernte Verwandtschaft. Die plastidäre FBP entstand durch Duplikation der cytosolischen Isoform, die SBP gruppiert zusammen mit eubakteriellen FBP-Sequenzen (Martin et al. 1996) und könnte einer weiteren Genduplikation der eukaryotischen FBP, die noch vor der Aufspaltung in cytosolische und plastidäre FBP stattfand, entstammen.

In dieser Arbeit wurden SBP und FBP-Sequenzen des Haptophyten *Prymnesium* und des Dinoflagellaten *Lingulodinium* etabliert. Weitere Sequenzen aus *Guillardia* (Cryptophyta) wurden im Rahmen einer Kooperation von der Arbeitsgruppe Uwe-G. Meier (Philipps-Universität Marburg) zur Verfügung gestellt. Es wurde somit die Präsenz der entsprechenden Gene, insbesondere der SBP, in allen photosynthetischen Abteilungen mit primären Plastiden (Rot-, Grünalgen und Glaucocystophyta) sowie komplexen Rhodoplasten (Haptophyta, Cryptophyta, Heterokontophyta, Dinoflagellaten) und komplexen Chloroplasten (Euglenophyta, Chlorarachniophyta) dokumentiert (siehe Abb. 21). Es ist daher wahrscheinlich, dass der eukaryotische Calvin-Zyklus inklusive der spezifischen SBP einmal früh in der primären photosynthetischen Linie etabliert und anschließend durch weitere Endosymbiosen verbreitet wurde. Die universelle Präsenz der SBP in Landpflanzen und Algen ist jedoch nur ein Teil einer komplexeren Geschichte, da die nicht-photosynthetischen

5. Diskussion B

Apicomplexa ebenso wie drei aplastidäre Abteilungen (Trypanosomen, Pilze, Ciliaten) auch ein SBP-Homolog besitzen. Die metabolische Funktion des entsprechenden Enzyms ist unklar, jedoch scheinen sie nicht als Ersatz einer FBP zu dienen, da diese heterotrophen Eukaryoten ebenfalls eine cytosolische FBP besitzen. Insbesondere die Präsenz der SBP in den Parasiten *Toxoplasma* und *Neospora* (Apicomplexa) ist erwähnenswert, da dieses Gen möglicherweise den ursprünglichen Calvin-Zyklus des Endosymbionten repräsentieren könnte, wenn es nicht über lateralen Gentransfer erworben wurde. Die phylogenetische Analyse (Abb. 21) unterstützt zudem einen Ursprung des Apicoplasten von Rotalgen (Foth und McFadden 2003; Coppin et al. 2005), der nach wie vor kontrovers diskutiert wird (Funes et al. 2002; Waller et al. 2003; Funes et al. 2003). Die SBP der aplastidären Trypanosomen (Euglenozoa) wurde kürzlich als Relikt eines gemeinsamen photosynthetischen Vorfahren aller Euglenozoa gedeutet (Hannaert et al. 2003), die vorliegende Analyse (Abb. 21) zeigt jedoch, dass die SBP von *Euglena* (sekundärer Chloroplast) mit den Sequenzen der Grünalgen nahe verwandt ist und keine Affinität zu trypanosomalen Genen besitzt. Trypanosomen haben demnach ihre SBP sehr wahrscheinlich unabhängig von *Euglena* erworben. Weiterhin beschränken sich die sehr schnell evolvierenden pilzlichen SBP-Sequenzen auf vier nahe verwandte Ascomyceten und sind mit Sicherheit kein Indikator für eine Photosynthese in Pilzen. Da in keinen anderen verfügbaren pilzlichen Genomen SBP-Gene existieren, ist es wahrscheinlicher, dass diese Ascomyceten die SBP später erworben haben, als dass es sich hierbei um ein ursprüngliches Gen der Pilze handelt, welches in allen anderen Pilzen verloren wurde.

Einige rezente Publikationen spekulieren auf der Basis der Chromalveolaten-Hypothese (Cavalier-Smith 1999) über eine photosynthetische Vergangenheit von Ciliaten und Oomyceten (Cavalier-Smith 1999; Patron et al. 2004; Fast et al. 2001). Die Präsenzen der SBP-Sequenzen von *Paramecium* und *Tetrahymena* sind potentielle Kandidaten für die Unterstützung dieses Postulats. Die phylogenetische Analyse (Abb. 21) belegt jedoch eindeutig, dass die SBP-Sequenzen von Ciliaten, Apicomplexa und Dinophyta (Alveolata) keinen gemeinsamen Ursprung haben. Da zudem kein Hinweis auf kryptische Plastiden in Ciliaten existiert, besitzen ihre SBPs sehr wahrscheinlich analog zu Ascomyceten und Trypanosomen andere Funktionen außerhalb einer Photosynthese und wurden unabhängig erworben.

Die Verteilung der SBP-Gene in den aplastidären heterotrophen Linien ließe sich zum einen durch (i) mehrere unabhängige laterale Gentransfers (LGT) erklären. Alternativ (ii) könnte es sich bei der SBP um ein „altes“ eukaryotisches Gen handeln, welches die Wirtszelle bereits vor der primären Endosymbiose besessen hat. In diesem Fall wäre die Sequenz

5. Diskussion B

Cyanidioschyzon (C) wahrscheinlich ein Relikt der ancestralen Genduplikation und diese divergente Kopie ging in allen anderen primär photosynthetischen Linien verloren. Dieses Szenario kommt zwar ohne LGTs aus, beinhaltet aber einige linienspezifische Genverluste. Unabhängig vom Szenario des Erwerbs demonstriert die phylogenetische Analyse der SBP (Abb. 21) jedoch letztlich, dass allein die Präsenz der SBP in heterotrophen Linien keinesfalls eine photosynthetische Vergangenheit voraussetzt.

Abgesehen von diesen „Besonderheiten“ der Genrekrutierung konnten in den vorliegenden Analysen auch einige diagnostische Gentransfers mit zumindest drei wichtigen evolutiven Aussagen identifiziert werden. (i) Im Bezug auf die SBP dokumentieren die Analysen einen gemeinsamen Ursprung von Viridiplantae, Rhodophyta und Glaucocystophyta. Die SBP wurde also nur einmal in einem gemeinsamen Vorfahren der primären photosynthetischen Linie rekrutiert und belegt somit auch den weithin akzeptierten monophyletischen Ursprung der Plantae (Rodríguez-Ezpeleta et al. 2005). (ii) Weiterhin haben alle plastidären FBPs (inkl. derer aus komplexen Algen) einen gemeinsamen Ursprung. In keinem der hier untersuchten Fälle erfolgte ein Ersatz durch eine duplizierte cytosolische Isoform, wie es beispielsweise bereits für die GapC-I gezeigt wurde (Liaud et al. 1997; Teich 2002). Möglicherweise hat eine Anpassung an die spezifische Regulation des plastidären Calvin-Zyklus über das Ferredoxin/Thioredoxin System (gezeigt für SBP, FBP, GAPDH, PRK in Landpflanzen; Buchanan 1980; Yano et al. 2001) einen solchen funktionellen Ersatz verhindert. (iii) Durch die im Rahmen dieser Arbeit isolierten SBP und FBP-Sequenzen aus den Abteilungen komplexer Algen mit Rhodoplasten (Haptophyta, Dinophyta, Cryptophyta, Heterokontophyta) wurde ein Datensatz bestehend aus allen photosynthetischen Eukaryotenlinien etabliert und die phylogenetischen Analysen belegen für die plastidären SBPs eine Monophylie der komplexen Algen mit Rhodoplasten (99-100%, siehe Abb. 21). Der phylogenetische Ursprung der SBP kann wahrscheinlich aufgrund der schnellen Evolutionsrate dieses Astes nicht aufgelöst werden. Eine gemeinsame Intronposition der Cryptomonade *Guillardia* und der Rotalge *Galdieria* (B) legt jedoch die Verwandtschaft zu Rotalgensequenzen nahe. Die plastidäre FBP liefert ein weiteres, wenn auch nicht so offensichtliches Beispiel für diese Monophylie. Unter der Annahme, dass *Euglena* und *Bigeloviella* ihre Gene durch einen lateralen Gentransfer erhalten haben, besitzen die Sequenzen von Diatomeen, *Emiliana*, *Guillardia* („Chromista“) und *Lingulodinium* einen gemeinsamen roten Ursprung. Analoge Beobachtung konnten im Rahmen der vorliegenden oder der früheren Diplomarbeit (Teich 2002) für die GAPDH (GapC-I; Teich 2002; Harper und Keeling 2003), Klasse II Aldolasen (FBA Klasse II; Patron et al. 2004) und die „grüne“ Phosphoribulokinase (s.o.; Petersen et al. 2006a) gemacht werden.

5. Diskussion B

5.6 Rückschlüsse der Einzelgen-Phylogenien auf die organismische Evolution komplexer Algen

5.6.1 Bedeutung der Einzelgenanalysen für die Entstehung komplexer Algen

Seit der Erkenntnis, dass ein Großteil eukaryotischer Algen durch sekundäre Endosymbiosen entstanden ist, beschäftigen sich Evolutionsbiologen mit der Frage nach dem Ursprung komplexer Plastiden. Ein umstrittenes Thema ist dabei die Anzahl unabhängiger Endosymbiose-Ereignisse, aus denen die komplexen Algen hervorgegangen sind. Akzeptiert ist mittlerweile, dass Chlorarachniophyta und Euglenophyta das Ergebnis zweier unabhängiger Endosymbiosen von Grünalgen sind (Gibbs 1978; Delwiche 1999; Archibald und Keeling 2002), die evolutive Verwandtschaft von Haptophyten, Diatomeen, Cryptomonaden und Dinophyten hingegen ist ungeklärt.

Grundlage für die phylogenetischen Analysen in der vorliegenden Arbeit bildeten Gene für Enzyme der katabolen Glykolyse und des anabolen Calvin-Zyklus, also des Primärstoffwechsels. Da die cytosolische Glykolyse bereits vor der plastidären Endosymbiose in der Wirtszelle existierte, sollten die Gene für diese cytosolischen Enzyme im einfachsten Fall die Wirtszelle repräsentieren. Enzyme, die spezifisch für den plastidären Calvin-Zyklus sind (z.B. PRK), sollten die Verwandtschaft der Plastiden klären lassen.

Die phylogenetischen Analysen von vier plastidären Enzymen (PRK, SBP, FBP, FBA Klasse II) lieferten in der vorliegenden Arbeit starke Hinweise auf die Monophylie komplexer Rhodoplasten. Die SBP konnte zwar auch in den nicht-photosynthetischen Ciliaten nachgewiesen werden, eine Unterstützung für eine ehemals photoautotrophe Lebensweise der Ciliaten bildet diese Analyse allerdings nicht (siehe Kap. 5.5). Eine signifikante wenn auch überraschende Entdeckung ist die „grüne“ PRK, die allen komplexen Algen mit Rhodoplasten gemein ist. Unabhängig davon, wie diese evolutiv erworben wurde (siehe Kapitel 5.4), liefert die Monophylie der Sequenzen dieser Algengruppe ein starkes Argument für den gemeinsamen Ursprung der Gene und weist auf eine gemeinsame organismische Herkunft hin. Die in der vorliegenden Arbeit beobachtete Verwandtschaft der Rhodoplasten komplexer Algen steht im Einklang mit den Ergebnissen phylogenetischer Analysen der GapC-I (Teich 2002; Harper und Keeling 2003). Diese Duplikation der cytosolischen GapC, ist spezifisch für komplexe Rotalgen (Teich 2002; Harper und Keeling 2003) und konnte in allen Algenabteilungen mit komplexen Rhodoplasten inklusive dem Apicomplexen *Toxoplasma* (Fast et al. 2001) nachgewiesen werden. Zusammen mit der vorliegenden Arbeit wurden

5. Diskussion B

somit insgesamt fünf nukleär kodierte Marker für den plastidären Calvin-Zyklus analysiert, die einen gemeinsamen Ursprung komplexer Rhodoplasten nahe legen.

In der 1999 verfassten Chromalveolaten-Hypothese (CA-Hypothese) postuliert Cavalier-Smith auf der Grundlage damals bekannter morphologischer (z.B. Flagellenstruktur), biochemischer (Chlorophyll c, Speicherkohlenhydrate) und molekularer Merkmale eine einzige sekundäre Endosymbiose, welche zur Entstehung der „Chromista“ (Hapto-, Crypto- und Heterokontophyta) und Alveolata führte. Die Basis für letztere Gruppe bilden 18S rDNA Analysen (Van de Peer und De Wachter 1997), welche auch durch Analysen mit konkatenierten nukleären Sequenzen bestätigt werden (Rodríguez-Ezpeleta et al. 2005). Alveolaten vereinen Dinophyta und die nicht-photosynthetischen Apicomplexa und Ciliaten. Relikte eines komplexen Plastiden stellt der von vier Membranen umgebene Apicoplast der parasitär lebenden Apicomplexa (inkl. Malariaerreger) dar. In Ciliaten allerdings wurde bisher kein Plastid oder Relikte desselben nachgewiesen, was im Widerspruch zur CA-Hypothese steht. Zwar konnte in der vorliegenden Arbeit ein gemeinsamer Ursprung der komplexen Rhodoplasten gezeigt werden, eine weitere notwendige Bedingung der Hypothese von Cavalier-Smith ist jedoch, dass auch die Wirtszellen aller als „Chromalveolata“ zusammengefassten Abteilungen monophyletisch sind, was in dieser Arbeit anhand der Gene für die cytosolische Glykolyse untersucht werden sollte.

Da die Glykolyse ein universeller Bestandteil des eukaryotischen Metabolismus ist, war es in der vorliegenden Arbeit möglich, die cytosolischen Enzyme von Ciliaten und Apicomplexa in die phylogenetischen Analysen mit einzubeziehen, was für Enzyme des plastidären Calvin-Zyklus aufgrund des Fehlens des Stoffwechselweges in diesen Organismengruppen nicht möglich ist. Die cytosolische Lokalisierung der Glykolyse spricht für eine Vererbung über die Wirtszelle und ihre Enzyme sollten im Kontext der CA-Hypothese monophyletisch sein. Durch keines der hier untersuchten Gene (TPI, ENO, cytosolische Aldolasen Klasse I und II, cytosolische FBP) konnte jedoch die Monophylie von Haptophyta, Cryptophyta und Heterokontophyta („Chromista“) und Dinophyta, Apicomplexa und Ciliaten (Alveolata) gezeigt werden. Diese Beobachtung wird auch von Analysen der 18S rDNA (Van de Peer und De Wachter 1997), cytosolischen Genen (Harper et al. 2005; Baldauf et al. 2000) und mitochondrialen Genen (Sánchez Puerta et al. 2004) bzw. den Analysen mitochondrialer Genome (Gray et al. 1998) geteilt. Die hinreichende Bedingung eines gemeinsamen Ursprungs der Wirtszellen ist dementsprechend nicht gegeben, was gegen die Chromalveolaten-Hypothese spricht.

5. Diskussion B

5.6.2 Die komplexe Evolution komplexer Algen - Tertiäre Endosymbiosen

Wie aber ist es zu erklären, dass phylogenetische Analysen von nukleär kodierten Genen für plastidäre Enzyme (diese Arbeit; Teich 2002; Harper und Keeling 2003; Patron et al. 2004; Kroth et al. 2005) und von plastidär kodierten Proteinen (Daugbjerg und Anderson 1997; Yoon et al. 2005) oft auf eine Monophylie der Sequenzen aus komplexen Algen mit Rhodoplasten hinweisen, die Analyse mitochondrialer (Gray et al. 1998; Sánchez Puerta et al. 2004) und cytosolischer Proteine aber nicht (diese Arbeit; Baldauf et al. 2000)?

Ein mögliches evolutives Szenario könnte mehrere parallele sekundäre Endosymbiosen, die zur Entstehung der komplexen roten Linie führten, beinhalten. Dies würde bedeuten, dass zwar die Wirtszellen aus unterschiedlichen Quellen stammten, die Endosymbionten jedoch identische bzw. sehr nah miteinander verwandte primäre Rotalgen waren. Demzufolge müssten alle Abteilungen komplexer Algen mit Rhodoplasten die „grüne“ PRK unabhängig voneinander erworben haben, was sehr unwahrscheinlich ist. Insbesondere im Falle der plastidären GapC-I setzt dies nach unabhängigen, parallelen Duplikationen der GapC mehrere parallele Transitpeptid-Rekrutierungen sowie die unabhängige Modifikation der Cosubstratspezifität voraus (siehe Teich 2002). Aufgrund der Komplexität dieser Voraussetzungen von PRK und GapC-I kann dieses Szenario ausgeschlossen werden.

Eine deutlich plausiblere Erklärung für die beobachteten Topologien würden tertiäre Endosymbiosen liefern, bei denen eine tertiäre Wirtszelle eine (bereits) komplexe Alge mit sekundärem Rhodoplasten zum Plastiden reduziert hat. Für einige Dinophyta wurde dies bereits beschrieben, wobei Haptophyta (*Karoldinium micrum*, *Karenia brevis*, *K. mikimotoi*; Tengs et al. 2000), Cryptophyta (*Dinophysis*; Hackett et al. 2003) bzw. Heterokontophyta (*Peridinium balticum*, *P. foliaceum*; Chesnick et al. 1997) als Endosymbiont fungieren. Im Bezug auf das Ensemble der Alveolata lässt sich mit diesem Szenario der Evolution auch erklären, warum Dinophyta, Apicomplexa und Ciliaten in phylogenetischen Analysen von 18S rDNA und anderen nukleär kodierten cytosolischen Proteinen eine signifikant unterstützte Gruppe bilden (Van de Peer und De Wachter 1997; Rodríguez-Ezpeleta et al. 2005), in den komplett sequenzierten Genomen von Ciliaten (*Paramecium*, *Tetrahymena*) aber bisher kein Hinweis auf eine plastidäre Vergangenheit gefunden wurde. Denn wahrscheinlich haben Ciliaten nie einen Plastiden besessen.

Aus den bisher verfügbaren fünf plastidären Markern lässt sich allerdings noch kein schlüssiges Szenario ableiten, welche Algenabteilungen einen sekundären bzw. tertiären Plastiden besitzen. Dementsprechend ist es notwendig, nach der Etablierung zusätzlicher Genome aus komplexen Algen mit Rhodoplasten (insbesondere aus Cryptomonaden und

5. Diskussion B

Dinophyten) weitere sorgfältige Gen-für-Gen Analysen durchzuführen, um die Evolution komplexer Algen endgültig zu entschlüsseln.

6. Zusammenfassung

In der vorliegenden Arbeit wurde anhand des GAPDH-Gensystems der Übergang von den Grünalgen zu terrestrischen Pflanzen untersucht, mit welchem auch die Duplikation der nukleär kodierten plastidären GapA korreliert. Die entstandene GapB erwarb eine charakteristische C-terminale Extension mit regulatorischer Funktion, die eine zweifelsfreie Identifizierung ermöglicht. Es konnten GapA und GapB-Sequenzen der einzelligen Grünalge *Mesostigma viride* etabliert werden, deren evolutive Stellung bisher höchst umstritten war. Die Präsenz der GapB und die phylogenetische Analyse beider Gene dokumentiert eindeutig, dass *Mesostigma* basal zu den Charophyceae und Landpflanzen abzweigt und so der primitivste Streptophyt ist. Damit erfolgte die *GapA/B*-Genduplikation bereits vor mehr als 700 Millionen Jahren. Darüber hinaus wurden die Intronpositionen von allen verfügbaren GAPDH-Sequenzen (*GapA*, *B*, *C* und *Cp*) aus *Mesostigma viride* und dem Lebermoos *Marchantia polymorpha* als Vertreter der primitivsten Landpflanzen ermittelt und mit denen von Referenzsequenzen verglichen. Abgesehen von gelegentlichen Intronverlusten ist die Genstruktur aller dieser GAPDH-Gene in Landpflanzen universell konserviert, was eine Intronstasis seit mindestens 500 Millionen Jahren belegt. Auch über die *GapA/B* und *GapC/Cp*-Genduplikation hinaus sind viele Intronpositionen konserviert und einige davon lassen sich sogar bis zu den Chlorophyceae zurückverfolgen, was auf eine konstante Vererbung seit einer knappen Milliarden Jahre hindeutet.

Der zweite Schwerpunkt dieser Arbeit beschäftigte sich mit so genannten komplexen Algen, die den Großteil des eukaryotischen Phytoplanktons der Weltmeere bilden und eine entscheidende ökologische Bedeutung besitzen. Ihre Plastiden entstanden durch sekundäre Endosymbiosen von Grün- und Rotalgen. Insbesondere die evolutiven Beziehungen der Haptophyta, Heterokontophyta, Cryptophyta und Dinophyta, die komplexe rote Plastiden besitzen, gab in der Vergangenheit Anlass zu kontroversen Diskussionen. Im Rahmen dieser Arbeit wurden anhand von nukleär kodierten Genen des Primärstoffwechsels die Verwandtschaftsbeziehungen dieser Algenabteilungen untersucht. Die Phylogenien mehrerer unabhängiger Gene für den plastidären Calvin-Zyklus wie etwa der Sedoheptulose-1,7-bisphosphatase und Fructose-1,6-bisphosphatase zeigen analog zur Glycerinaldehyd-3-phosphat Dehydrogenase eindeutig die Monophylie der entsprechenden Sequenzen aller Algenabteilungen mit komplexen Plastiden, weshalb sie in der Literatur auch als „lucky genes“ bezeichnet wurden. Zusammen mit der überraschenden Entdeckung einer „grünen“ Phosphoribulokinase in diesen Algenabteilungen, welche ein einziges Mal von einer

6. Zusammenfassung

Grünalge rekrutiert wurde, legen diese unabhängigen molekularen Marker einen gemeinsamen Ursprung der komplexen roten Plastiden durch eine einzige sekundäre Endosymbiose nahe. Im Gegensatz dazu lassen die phylogenetischen Analysen der Gene für die cytosolische Glykolyse keine gemeinsame evolutive Vergangenheit der Wirtszellen von Algen mit komplexen Rhodoplasten erkennen. Dies steht im Widerspruch zur Chromalveolaten-Hypothese, die einen gemeinsamen Ursprung dieser Algenabteilungen postuliert, was auch eine Monophylie der Wirtszellen voraussetzt. Ein evolutives Szenario, welches die eindeutigen Daten plastidärer Marker mit denen der Wirtszelle plausibel in Einklang bringt, beinhaltet tertiäre Endosymbiosen, bei denen eine Alge mit sekundär erworbenem Plastiden von einer eukaryotischen Wirtszelle aufgenommen und zum tertiären Plastiden reduziert wurde.

7. Material und Methoden

7.1 Herkunft des Pflanzenmaterials, der DNA bzw. RNA

***Pavlova lutheri* (Haptophyta; SAG 926-1)**

Der *Pavlova lutheri* Stamm 926-1 der Sammlung von Algenkulturen der Universität Göttingen (SAG) wurde bei 22°C (Kulturraum der AG Mendel) und einem Tag/Nacht-Rhythmus von 14,5 h Licht/9,5 h Dunkelheit in Brackish-Water Medium (SAG; siehe <http://www.epsag.uni-goettingen.de/html/culturemedia.html>) unter ständigem Rühren bzw. Schütteln kultiviert. RNA und genomische DNA für PCR-Experimente wurden nach dem Standardprotokoll (siehe 7.11.2) isoliert.

***Prymnesium parvum* (Haptophyta; SAG 127.79)**

Prymnesium parvum entstammt der Sammlung von Algenkulturen der Universität Göttingen (SAG; Stamm 127.79). Die Kulturbedingungen sind zu denen von *Pavlova lutheri* identisch.

***Hanusia phi* (Cryptophyta)**

Die für PCR-Experimente benötigte cDNA wurde durch Phagengroßkultur einer im Labor vorliegenden λ NM1149 cDNA-Bank gewonnen.

***Guillardia theta* (Cryptophyta)**

Die genomische DNA wurde von der Arbeitsgruppe Uwe-G. Maier (Philipps-Universität Marburg) zur Verfügung gestellt.

***Pyrocystis lunula* (Dinophyta)**

Eine λ ZAPII cDNA-Bank des Dinophyten *Pyrocystis lunula* wurde von J. Woodland Hastings (Harvard University Cambridge) zur Verfügung gestellt. Für PCR-Experimente wurde die cDNA durch Massenexzision isoliert.

***Lingulodinium polyedrum* (früher: *Gonyaulax polyedra*, Dinophyta)**

Eine λ ZAPII cDNA-Bank des Dinophyten *Lingulodinium polyedrum* wurde von J. Woodland Hastings (Harvard University Cambridge) zur Verfügung gestellt. Für PCR-Experimente wurde die cDNA durch Massenexzision isoliert.

7. Material und Methoden

***Phaeodactylum tricornutum* (Diatomee, Heterokontophyta)**

Die genomische DNA von *Phaeodactylum tricornutum* wurde von M.F. Liaud (TU-Braunschweig) zur Verfügung gestellt.

***Bigelowiella natans* (Chlorarachniophyta)**

Genomische DNA aus *Bigelowiella natans* wurde von G. McFadden (Melbourne) zur Verfügung gestellt.

***Euglena gracilis* (Euglenophyta)**

Die genomische DNA von *Euglena gracilis* wurde von W. Martin (Heinrich-Heine Universität Düsseldorf) zur Verfügung gestellt.

***Chondrus crispus* (Rhodophyta)**

Das Pflanzenmaterial von *Chondrus crispus* stammt aus dem Felswatt von Helgoland. DNA und RNA wurden mittels der Standardmethoden isoliert.

***Marchantia polymorpha* (Lebermoos)**

RNA und genomische DNA des Lebermooses *Marchantia polymorpha* lagen zu Beginn dieser Arbeit bereits aufgereinigt vor (siehe Petersen 1999).

***Chara vulgaris* (Charophyceae)**

RNA und DNA von *Chara vulgaris* lagen bereits im Labor aufgereinigt vor (siehe Petersen 1999).

***Coleochaete scutata* (Charophyceae; B 3.90)**

Die genomische DNA von *Coleochaete scutata* (SAG-Stamm B 3.90) lag bereits aufgereinigt im Labor vor (siehe Petersen 1999).

***Chlorokybus atmophyticus* (Charophyceae)**

Die genomische DNA von *Chlorokybus atmophyticus* lag bereits im Labor vor (siehe Petersen 1995).

7. Material und Methoden

***Mesostigma viride* (Streptophyta)**

Pflanzenmaterial der Grünalge *Mesostigma viride* wurde freundlicherweise von B. Becker (Universität Köln) zur Verfügung gestellt. Genomische DNA sowie Gesamt-RNA wurden aus 1,2 g Algenmaterial nach dem Standardprotokoll isoliert. Dabei wurden zur Polysaccharidfällung 0,15Vol EtOH (100%) eingesetzt. (RNA c = 1,332 µg/µl; DNA c = 80 ng/µl)

***Cladophora rupestris* (Ulvophyceae)**

Die genomische DNA von *Cladophora rupestris* lag bereits im Labor vor (siehe Petersen 1995).

7.2 *Escherichia coli* Stämme

POP13	Schwarz-Sommer et al. 1987
K803	Wood 1966
KW251	Promega
SOLR	Stratagene
XL1 Blue MRF'	Stratagene

7.3 Plasmid- und Bakteriophagen Vektoren

λEMBL3	Frishauf et al. 1983; (Promega)
λZAPII	Short et al. 1988; (Stratagene)
λNM1149	Murray 1983
pBluescript II SK+	Stratagene
pCR 2.1	Invitrogen

7. Material und Methoden

7.4 Enzyme

Restriktionsendonukleasen

Restriktionsendonukleasen wurden von der Firma MBI Fermentas bezogen.

Sonstige Enzyme

Alkalische Phosphatase CIAP	MBI Fermentas, # EF0341
Bio-X-Act Long Polymerase	Bioline, # BIO-21049
Biolase Diamond	Bioline, # BIO-21B59
DNase	Boehringer Mannheim, # 104159
Klenow-Polymerase	MBI Fermentas, # EP0052
Proteinase K	Merck, # 124568.0500
RNase	Serva, # 34390.03
RNase H	MBI Fermentas, # EN 0201
Superscript III Reverse Transkriptase	Invitrogen, # 18080-044
Thermoscript Reverse Transkriptase	Invitrogen, # 11146-024
T4 DNA-Ligase	MBI Fermentas, # EL0011
T4 Polynukleotid Kinase	MBI Fermentas, # EK0031
Taq DNA Polymerase	Qiagen, # 1064299
Taq DNA Polymerase (rekombinant)	Invitrogen, # 10342-020

7.5 Reagenzsätze (Kits)

Geneclean Turbo (Q-BIOGENE, # 1102-200)
DNase Quick Clean (Bioline, BIO-37042)
ThermoScript™ RT-PCR System (Invitrogen, # 11146-024)
Kit für radioaktive Sequenzierung (Pharmacia # 27-1674-01)
ALF express Auto Read Sequencing Kit (Amersham # 27-2690-02)
BigDye Terminator v3.1 Sequencing Standard Kit (Applied Biosystems, # 4336935)
Kit für Maxi-Präparation (Genomed # 220020)
Lambda EMBL3 BamHI Vector Kit (Stratagene # 241211)
ZAP-cDNA Synthesis Kit (Stratagene # 200400)

7. Material und Methoden

7.6 Sonstiges

λ /HindIII-Marker	MBI Fermentas, # SM0101
1 kb-Ladder	Invitrogen, # 15615-016
100bp DNA Ladder (O'RangeRuler)	MBI Fermentas, # SM0623
BSA	MBI Fermentas, # B14
p(dN)6 Hexanukleotide	Pharmacia, # 27-2166-01

7.7 Chemikalien

Die verwendeten Chemikalien stammen von den Firmen Roth, Merck und Sigma. Radiochemikalien lieferte die Firma Hartmann Analytics, Braunschweig. Die Röntgenfilme (X-OMAT AR) sowie die Entwickler- und Fixierlösung stammten von der Firma Kodak.

7.8 Medien für *E. coli* und Lambda-Phagen

CY:	Casamino Acids	10 g/l
	Yeast-Extract	5 g/l
	NaCl	3 g/l
	KCl	2 g/l
	Tris-HCl 1 M, pH 7,5	25 ml
	MgSO ₄ 1 M	5 ml
LB:	Bacto-Trypton	10 g/l
	Yeast-Extract	5 g/l
	NaCl	10 g/l
	pH 7,5 einstellen	
	(für Platten) Agar	15 g/l)
LB-AIX:	Ampicillin	100 mg/l
	IPTG	47,5 mg/l
	X-Gal	40 mg/l

7. Material und Methoden

dyt:	Bactotrypton	16 g
	Yeast Extract	10 g
	NaCl	5 g
NZ:	NZ-Amine	10 g/l
	NaCl	5 g/l
	Casamino Acids	1 g/l
	MgSO ₄ •7H ₂ O	2 g/l
	pH 7,5 einstellen	
	Agar	15 g/l
TOP-Agar:	Agarose	0,7 % (w/v)
	MgSO ₄	10 mM
SM-Puffer:	NaCl 5 M	2 % (v/v)
	MgSO ₄	0,1 % (w/v)
	Tris-HCl pH 7,5 1 M	5 % (v/v)
	Gelatine 2 % (w/v)	0,5 % (v/v)

7.9 Geräte

Automatische Sequenzierer	ABI Prism 310 Genetic Analyzer; Pharmacia LKB A.L.F. DNA Sequenzer mit Compaq DeskPro 386/20e
Computersysteme	MacIntosh G3 Windows PC
Elektrophoresekammern	BIO RAD, Sub-Cell GT
Elektrophorese-Netzgeräte	Ingenieurbüro Fischer für Agarosegele Pharmacia ECPS 3000/150 für Sequenzgele LKB 2297 Macrodrive 5 für Sequenzgele
Elektroporator	BIO-RAD Gene Pulser II und BIO-RAD Pulse Controller II
Elektroporationsküvetten	VWR, # 732 1802

7. Material und Methoden

Kühlwasserbad	Julabo F 10-Kryostat
Kühlzentrifugen	Sorvall RC5B mit Rotoren: GS3, GSA, SS 34, SA-600
PCR-Maschine	Eppendorf Mastercycler
Pipetten	Gilson 2, 20, 200, 1000
Thermostat	Eppendorf 5320
Tischzentrifugen	Heraeus, Biofuge fresco und Biofuge 22R
Ultrazentrifuge	Beckmann TL100 mit Rotor: Ti 70.1 und SW 41
Untertischzentrifuge	Beckmann, GPKR
Vakuumzentrifuge	Bachofer
Wasserbäder	GFL

7.10 Arbeiten mit *E. coli*

Die Anzucht von *E. coli* auf Platten, in Flüssigmedien, die Transformation CaCl_2 -kompetenter *E. coli* Zellen mit Plasmid-DNA, Transformation durch Elektroporation und die Transfektion mit λ -DNA erfolgte nach Standardprotokollen (Sambrook et al. 1989). Zur Anzucht von *E. coli* Kulturen wurden folgende Medien verwendet: CY-Medium, LB-Medium, NZ-Medium und dyt-Medium.

7.11 Arbeiten mit Nukleinsäuren

7.11.1 Standardmethoden

Allgemeine Techniken zum Arbeiten mit Nukleinsäuren wurden nach den Standardprotokollen nach Sambrook et al. (1989) durchgeführt. Hierzu zählen:

Methoden zur Isolierung von Nukleinsäuren (Plasmid-Minipräparationen aus *E. coli* Zellen, präparative Isolierung von Plasmiden aus *E. coli* Zellen, Isolierung von λ -Phagen-DNA).

Methoden zur Reinigung von Nukleinsäuren (Phenolisierung, Ethanolpräzipitation, Isopropanolpräzipitation, Sucrose- und CsCl -Dichtegradientenzentrifugation, Sephadex-G-50-Chromatographie, Agarose-Gelelektrophorese, Konzentrationsbestimmung von Nukleinsäuren).

7. Material und Methoden

Enzymatische Modifikationen von Nukleinsäuren (Restriktionsschnitte, Ligation mit T4-DNA-Ligase, Auffüllreaktion mit Klenow-Polymerase, Phosphorylierung von DNA mit T4-Kinase, Dephosphorylierung von DNA mit alkalischer Phosphatase).

Methoden zum Transfer von Nukleinsäuren auf Membranen (Southernblot, Transfer von Phagen-Plaques und Bakterienkolonien auf Nitrozellulose-Filter).

Nachweis von Nukleinsäuren (Konzentrationsmessungen von RNA und DNA, Herstellung von radioaktiv und nicht-radioaktiv markierten Sonden, radioaktive und nicht-radioaktive Hybridisierung mit DNA bzw. Oligonukleotidsonden).

7.11.2 RNA- und DNA-Isolation

Isolierung von Gesamt-RNA

100ml Extraktionspuffer auf 65° C in einem 250 ml Zentrifugenbecher vorwärmen. 20 g Pflanzenmaterial in flüssigem Stickstoff möglichst fein mörsern. Pulvertrichter und Löffel in Stickstoff vorkühlen. 200 µl β-Mercaptoethanol zum Extraktionspuffer zufügen. Stickstoff im Mörser verdampfen lassen. Das trockene Material möglichst schnell in den vorgewärmten Extraktionspuffer geben und kräftig schütteln. Zum Entfernen der Zelltrümmer im GSA-Rotor für 5 min bei 4° C und 4.000 g zentrifugieren. Der Überstand wird in einen neuen Zentrifugenbecher überführt und mit 1 Volumen Phenol/Chloroform/Isoamylalkohol (25:24:1) versetzt. 1 min extrahieren. Für 10 min bei 4° C und 6.000 g zentrifugieren. Zur klaren Oberphase (sonst Phenolisierung wiederholen) langsam 0,15 Volumen Ethanol zugeben, um selektiv Polysaccharide zu fällen und anschließend mit 1 Volumen Chloroform/Isoamylalkohol (24:1) 1 min extrahieren. Für 10 min bei 4° C und 6.000 g zentrifugieren. Beim Abnehmen der Oberphase keine Interphase mit überführen. 0,25 Vol 10 M LiCl zufügen und 2 – 3 h oder ü. N. im Eisbad im Kühlraum die RNA fällen. Für 30 min bei 4° C und 6.000 g zentrifugieren. Den Überstand aufbewahren, da er die genomische DNA enthält. Das Pellet in 2 x 2,5 ml SSTE aufnehmen, welcher auf 50° C vorgewärmt ist und in ein Nalgene-Röhrchen überführen, 2 Volumen Ethanol zugeben, 1 – 2 h bei -20° C fällen. Für 40 min bei 4° C und 40.000 g im SS-34 Rotor zentrifugieren. Überstand verwerfen, Pellet in 2 – 3 ml H₂O aufnehmen. 5 µl der Gesamt-RNA für ein Testgel aliquotieren, den Rest bei -80° C lagern.

7. Material und Methoden

Extraktionspuffer (100ml)	2 % CTAB (w/v)	2 g
	100 mM Tris-HCl pH 8,0	10 ml 1 M Tris-HCl pH 8,0
	25 mM EDTA pH 8,0	5 ml 0,5 M EDTA pH 8,0
	2 M NaCl	9,68 g
	8 % Polyclar AT (w/v)	8 g
	autoklavieren	
	0,2 % (v/v) 0,14 M β -Mercaptoethanol kurz vor Gebrauch zugeben	

SSTE (100 ml):	1 M NaCl	20 ml 5 M NaCl
	0,5 % SDS	5 ml 10 % SDS
	10 mM Tris-HCl pH 8,0	1 ml 1 M Tris-HCl pH 8,0
	1 mM EDTA pH 8,0	0,2 ml 0,5 M EDTA pH 8,0
	autoklavieren	

Isolation von genomischer DNA

Den bei der RNA-Fällung mittels Lithiumchlorid anfallenden Überstand mit 1/10 Vol 3 M NaOAc pH 6,5 und 2 Vol 100 % Ethanol versetzen. 1 h bei -20° C fällen lassen. Für 20 min bei 4° C mit 13.000 rpm zentrifugieren. Das Pellet mit 70 % Ethanol waschen. Für 10 min bei 4° C mit 13.000 rpm zentrifugieren. Überstand dekantieren, anzentrifugieren, restliches Ethanol entfernen. 5 min bei RT trocknen lassen. Genau 10 ml TE zugeben und 1-2 h leicht schütteln lassen. Sollte die genomische DNA zu trocken geworden sein, so kann es mehrere Tage dauern, bis sie sich wieder löst.

Poly(A)+-RNA-Gewinnung

40 mg Oligo-dT-Cellulose werden in ein steriles 50 ml Falcon-Röhrchen eingewogen, in 25 ml Bindungspuffer suspendiert, 1 Minute bei 3.400 rpm zentrifugiert (Beckman Untertischzentrifuge) und nach Dekantieren des Überstandes in 10 ml Bindungspuffer aufgenommen. Die isolierte Gesamt-RNA hinzugeben und 30 min bei RT schütteln. Für 1 Minute bei 3.400 rpm zentrifugieren (Untertischzentrifuge). Überstand abgießen, aber noch aufbewahren. Dreimaliges Waschen der Cellulose mit 25 ml Bindungspuffer (resuspendieren), 1 min bei 4.000 rpm zentrifugieren (Untertischzentrifuge). Eine Glassäule mit 25 ml Waschpuffer spülen. Oligo-dT-Cellulose mit 25 ml Waschpuffer waschen und in 10 ml Waschpuffer auf die Säule gegeben. Mit 2 x 900 ml Waschpuffer über eine Drucksäule waschen. Die mRNA wird in 3 Fraktionen à 3 ml durch auf 55° C vorgewärmten Elutionspuffer eluiert und in 15 ml Corex-Röhrchen (silanisiert) aufgenommen. Zugabe von

7. Material und Methoden

300 µl 3 M NaOAc (pH 6,5) und 7 ml EtOH, mischen, verschließen des Röhrchens mit Parafilm, Fällern der mRNA ü. N bei -20° C. Pelletieren durch Zentrifugation für mindestens 1 h bei 12.000 g. Das Pellet mit 1 ml 70% Ethanol waschen, 30 min zentrifugieren (12.000 g). Überstand dekantieren, Trocknen des Pellets für 30 min. Die Poly(A)⁺ wird in 20 µl H₂O aufgenommen und bei -80° C gelagert.

Bindungspuffer (400 ml):	400 mM NaCl	32 ml 5 M NaCl
	10 mM Tris-HCl pH 7,5	4 ml 1 M Tris-HCl pH 7,5
	0,2 % SDS	8 ml 10 % SDS
	H ₂ O	365 ml
	autoklavieren	
Waschpuffer (900 ml):	100 mM NaCl	18 ml 5 M NaCl
	10 mM Tris-HCl pH 7,5	9 ml 1 M Tris-HCl pH 7,5
	0,2 % SDS	18 ml 10 % SDS
	H ₂ O	855 ml
	autoklavieren	
Elutionspuffer (100 ml):	10 mM Tris-HCl pH 7,5	1 ml 1 M Tris-HCl pH 7,5
	H ₂ O	99 ml
	autoklavieren	

7.11.3 *in vivo* Exzision rekombinanter λZAPII-Phagen

Die *in vivo* Exzision (Einzelklon und Massenexzision) von rekombinanten λZAP cDNA Klonen wurden mittels des ExAssist® Helfer Phagen der Firma Stratagene (# 211203) laut Protokoll durchgeführt. Anstelle des *E. coli* Stammes XL0LR wurden jedoch *E. coli* SOLR verwendet.

7. Material und Methoden

7.11.4 PCR-Techniken

7.11.4.1 RACE (rapid amplification of cDNA ends)

Die Technik der RACE-PCR wurde von Frohman 1993 beschrieben.

7.11.4.2 Reverse Transkriptase PCR (RT-PCR)

In dieser Arbeit wurden für RT-PCR Reaktionen „SuperScriptIII Reverse Transcriptase“ und das „ThermoScript™ RT-PCR-System“ (beide Invitrogen) verwendet. Die Arbeiten mit den Enzymen erfolgten nach den Protokollen der Firma. Die ThermoScript™ erlaubt aufgrund der Flexibilität des Enzyms Modifikationen der Reaktionstemperaturen und –zeiten. Es wurde folgendes Programm für die cDNA-Synthese mittels ThermoScript™ RT-PCR-System erstellt:

20 min 50°C

20 min 55°C

20 min 60°C

20 min 65°C

7.11.4.3 PCR mit genomischer DNA

Für PCR Reaktionen mit genomischer DNA wurde die Bio-X-Act Long Polymerase (Bioline, # BIO-21049) laut Angaben des Herstellers verwendet. Die Amplifikation erfolgte durch eine so genannte nested-PCR. Dabei wurden die äußeren (ersten) Primerpaare in einer Konzentration von 10 pmol/50 µl Ansatz eingesetzt, die inneren (zweiten) Primerpaare in einer Konzentration von 100 pmol/50 µl Ansatz.

7.11.5 Sequenzierung

Sequenzierung mit dem Pharmacia ALF

Zur Sequenzierung mit dem ALF wurde das ALF express Auto Read Sequencing Kit (Amersham # 27-2690-02), wie im Protokoll beschrieben, verwendet.

7. Material und Methoden

ABI-Kapillar-Sequenzierung

Sequenzierung mit dem ABI Prism 310 Genetic Analyzer

Für Sequenzierungen im Haus wurde der ABI Prism 310 Genetic Analyzer von der AG Jahn (Institut für Mikrobiologie, Braunschweig) zur Verfügung gestellt. Es wurde das BigDye Terminator v3.1 Sequencing Standard Kit (Applied Biosystems, # 4336935) nach dem Standard Protokoll verwendet.

Auftragssequenzierung

Die Auftragssequenzierung erfolgte durch die Firma Sequence Laboratories Göttingen GmbH (SEQLAB).

Adresse: Sequence Laboratories Göttingen GmbH (SEQLAB)
 Postfach 3343
 37023 Göttingen

7.12 *in silico* Arbeiten

7.12.1 Datenbanken

NCBI – National Center for Biotechnology Information

<http://www.ncbi.nlm.nih.gov/>

JGI – DOE Joint Genome Institute

<http://genome.jgi-psf.org>

TBestDB- Taxonomically Broad EST Database (früher PEPdb)

<http://amoebidia.bcm.umontreal.ca/pepdb/searches/login.php>

Kazusa DNA Research Institute:

<http://www.kazusa.or.jp/en/plant/porphyra/EST/index.html>

Cyanidioschyzon merolae Genome Project:

<http://merolae.biol.s.u-tokyo.ac.jp/>

TIGR – The Institute for Genomic Research:

<http://www.tigr.org>

PlasmoDB – The Plasmodium Genome Resource:

<http://www.plasmodb.org/plasmo/home.jsp>

7. Material und Methoden

7.12.2 BLAST-Programme

Für die Etablierung von Sequenzen aus Datenbanken wurden die jeweils verfügbaren BLAST-Programme verwendet:

blastn: Nukleotidsequenz gegen Nukleotid Datenbank

blastx: übersetzte Nukleotidsequenz gegen Protein Datenbank

tblastn: Proteinsequenz gegen übersetzte Nukleotid Datenbank

tblastx: übersetzte Nukleotidsequenz gegen übersetzte Nukleotid Datenbank

7.12.3 Vorhersageprogramme für die intrazelluläre Lokalisierung von Proteinen

PSORT (Protein Sorting Server): <http://psort.ims.u-tokyo.ac.jp/>

CBS Prediction Servers: <http://www.cbs.dtu.dk/services/>

SignalP – Signalsequenzen in eukaryotischen Sequenzen

ChloroP – Transitpeptide für Chloroplasten in Pflanzen

TargetP – subzelluläre Lokalisation von Proteinen (u.a. Mitochondrium, Chloroplast)

7.12.4 Phylogenie-Programme

Programme für die Erstellung und Bearbeitung phylogenetischer Analysen sind auf folgender Internetseite erhältlich: <http://evolution.genetics.washington.edu/phylip/software.html> .

ClustalX (Version 1.81) stammt von: <ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/> .

7.12.5 Phylogenetische Analysen

Für die phylogenetischen Analysen in der vorliegenden Arbeit wurde, soweit nicht anders beschrieben, die Treefinder Version vom Oktober 2005 mit der WAG-Matrix (Whelan und Goldman 2001) verwendet. Die statistische Unterstützung der Genstammbäume wurde über 100 Bootstrap Analysen (Bootstrap Properties; BP) und 100 Edge Support (ES) Analysen bestimmt.

7. Material und Methoden

7.13 Primerliste

Nr.	Organismus/Gen	Sequenz
001	Chlorokybus GapA	TGATGAGCACCTTCTTGG
004	Chara GapA	AGAAACCCGCTCAACCTG
005	Chara GapA	CGGAGAGCGATACCATTG
018	Oligo dG-Primer	AAGAATTCGGGIIGGGIIGGGIIG
023	Oligo dT18 (NotI/EcoRI-Adapt.)	CCGCGGCCGCGAATTCTTTTTTTTTTTTTTTTTTTT
028	pBS-rev (Gille)	GGAAACAGCT ATGACCATGA TTAC
029	pBS-uni (Gille)	CGTAATACGA CTCACTATAG GGC
030	PYK-for(+)	GAYACNAAYGGNCCNGARAT
031	PYK-rev(-)	ARRACNCCNCKNGCNACCAT
032	FBA-I-for(+) ₁	ATHYTNTTYGARGARACNYTNTAYCA
033	FBA-I-for(+) ₂	GTNCCNATHGTNGARCCNGA
034	FBA-I-rev(-) ₁	AYYTCNGGYTCNACDATNGGNAC
035	FBA-I-rev(-) ₂	RTTNATNGCRTTNARRT
036	FBA-II-for(+)	GYNATGCAYCARGAYCAYGG
037	FBA-II-rev(-)	TCNGTRTCDATRTTNACYTT
038	FBP-for(+)	GTNGTNTTYGAYCCNYTNGAYGG
039	FBP-rev(-)	SCYTKNCCNCCNGCYTGyTC
049	Oligo dT20	TTTTTTTTTTTTTTTTTTTTTTT
051	GAPDH (WYDNE-;Jörn)	CCAYTCRTTTRTCRTACCA
052	GAPDH (GINGF+;Jörn)	GSNATHAAYGGNTTYGG
053	GAPDH-Reverse(Gille)	CTGTADCCCMAYTCRTTGTCTRTACCA
054	GAPDH-Foward(Gille)	AAGRtyGGHATCAAYGGHTTYGG
059	PGK:for +(147)	AACGACGCNTTYGGNACNGCNCA
060	PGK:rev-(384)	TCSAGSSWNGCNCCNCCNCCNGT
082	PYK-for(G/C)	GCSTSGACACSAARGGNCCNGARAT
083	PYK-rev(G/C)	ATSCCSAGGTCSCCNCKNGCNACCAT
084	TPI-for(GC)	TTCTTCGTSGGSGGNAAYTGGA
085	TPI-rev(GC)	ATSGCCASACSGGYTCRTANGC
086	PGK-for(GC)	AACGACGCSTTCGGSACSGCNCA
087	PGK-rev(GC)	TCGAASACSCCSAYSGGSCCRTTCCA
088	FBAI-for(GC)	GGSGYSATCCTSTTCSANGARAC
089	FBAI-rev(GC)	AYCTCSGGCTCSACGATSGGNAC
090	FBP-for(GC)	GTSGTSTTCGACCCSCTNGAYGG
091	FBP-rev(GC)	GCCTSSCCSCCSGCTGYTC
092	ENO-for(GC)	GACTCSCGSGGSAACCSACSGTNGA
093	ENO-rev(GC)	GGSGCGAASCCSCCCTCGTCNCCNAC
094	ENO-for(GC)	GTSGGSGACGAGGGSGGSTTCGCSCC
095	ENO-rev(GC)	CASGGSGCSCCSGTCTTDTATYTG
096	G6PI-pla-for(GC)	ACSGTSTACGSAAYAARGG
097	G6PI-pla-rev(GC)	TCSACSCCSGGCTGGTGGTANGC
098	PGM-for(GC)	ATGGGSAACWSSGAGGASGGNCA
099	PGM-for(GC)	CACGTSACSTWCTTCTKSAAYGG
100	PGM-rev(GC)	CCSGTGTGSCCSACCATRTC
101	G6PDH-for(GC)	ATCGACCACTACCTSGGNAARGA
102	G6PDH-rev(GC)	AGSACCTTSACCTTYTCRT
103	PYK-for(AT)	ATGTTWGATACWAARGGNCCNGARAT
104	PYK-rev(AT)	CCTAAATCWCCWCKNGCNACCAT
105	G6PI-for(AT)	GCWTTTTGGGATTGGGTNGGNGG
106	G6PI-rev(AT)	CCWAATTCWACWCCCCAYTRRTCRAA

7. Material und Methoden

109	6PGDH-for(GC)	CTSGCSGGSGCTSGCSGTNATGGG
110	6PGDH-rev(GC)	TCGATSCCGTTGTGSACCATYTTNAC
111	PFK-PPI-for(GC)	CAGGCSCCSGGSGGNCAAYAGT
112	PFK-PPI-rev(GC)	SCCTSGTASCCRAAYTCRTG
113	PFK-for(GC)	ACSTGCGSGSGSCTSTGYCCNGG
114	PFK-rev(GC)	TCGTTGTGCGATSGTYTTNGG
117	RPE-for(GC)	TTCGTSCCSAACATCACNATHGG
118	RPE-rev(GC)	CCSCCATCSACCTCDATCCANGG
119	Chondrus PGM	TTCGGTATCGCCCAAACAATG
124	SBP-rev(GC)	ACSGGSACCATSCCSCCSGTRTA
125	PRK-for(GC)	ACSGTSATCTGCCTSGACGAYTAYC
126	PRK-rev(GC)	TCSGCCATGTCSCGCTGGATYTTCC
127	RPI-for(GC)	ATCGACGGSGCSGACGTGGTNG
128	RPI-rev(GC)	TAGTTSCCGTTGTCSGTNAC
129	TKL-for(GC)	ATCGTSGACMACTACACNTA
130	TKL-rev(GC)	TGGTSGTSGGSCCGTCCTCNCC
146	Pavlova PRK	AGATGTCTGAAGGNAGATGG
165	Chondrus PRK	CCCATCTACAACCACGAG
166	Chondrus PRK	CGCCGGTCTCGTGGTTGT
167	Chondrus SBP (14)	TCGGCGTGTGGAAAGGTG
168	Chondrus SBP (14)	CCGACGGAAAAGTTAGTA
169	Chondrus SBP (16)	CTGTGCGGCACCATCTTCG
170	Chondrus SBP (16)	CGAAGATGGTGCCGACAG
182	Pavlova PRK	CGACCTGTCCATCTACCT
183	Pavlova ENO	ATGCCGTGCTTCAACCTT
187	Pyrocystis PRK (10+12)	ACCACGACACGGGCTTCA
188	Pyrocystis PRK (10+12)	CATAGATGGGGTGGAGGC
189	Gonyaulax PRK (21)	GCACCCAATCTACGACAA
190	Gonyaulax PRK (21)	TTGTCTGTAGATTGGGTGC
193	GAPDH-GINGF(-)	ATNCKNCCRAANCCRTT
194	GAPDH-GINGF(+)#NotI	AAFFAAAAAFCFFCCFCFSNATHAAYFFNTTYGG
195	GAPDH-GINGF(+)	GSNATHAAYGGNTTYGG
196	GAPDH-GINGF(+)#EcoRI	CCGAATTCTGSNATHAAYGGNTTYGG
197	GAPDH-CLAP(+)	TGYACNACNAAYTGYHT
198	GAPDH-CLAP(-)	ADRCARTTNGTNGTRCA
199	GAPDH-WYDNE(-)#NotI	AAGGAAAAAAGCGGCCGCCAYTCRTTTRTCRTACCA
200	GAPDH-WYDNE(-)	CCAYTCRTTTRTCRTACCA
200	GAPDH-WYDNE(-)	CCAYTCRTTTRTCRTACCA
201	GAPDH-WYDNE(-)#EcoRI	CCGAATTCAYTCRTTTRTCRTACCA
202	GAPDH-WYDNE(+)	TGGTAYGAYAAYGA
212	Euglena SBP (13)	GTCCGCGCTGAGGGTGGC
213	Euglena FBP (21)	TTACAGGGATGGCAGATG
214	Euglena FBP (28)	CCGTGCGCCCAATCAAT
215	Chondrus 6PGDH (15)	CGCCGCCTCGCCGCACCC
216	Porphyra FBPpla	CCCAGTCGTCAACACCAT
217	Chondrus G6PDH (9)	ACCAATTCGCATAGAGGA
218	Euglena FBP (28)	CGGAAAGGTTGGTTGCTG
228	Pavlova FBA (06-F02)	CGCTACGCGGTCATCTGC
229	Pavlova PRK	CTTGCCGGTCGCCTTCTCG
230	Pavlova ENO	CGGCGACGACATCTTCTGC
231	Pavlova TPI	CGCGATCACAACCGTCTC
232	Pyrocystis PYK	AGGGCAGCATCATCTCT
233	Porphyra GPI	ACGGGGTGTCGCTGGATT

7. Material und Methoden

234	Euglena PRK	GTTTGACCGCCCTGGACC
235	Euglena PRK	GTCAAAATCTGGCTTCCT
236	Euglena SBP (15)	CCGTCGGCACCATCTTCG
237	Chondrus 6PGDH (8)	AGAGCATCGCCGCCAAGA
239	Euglena SBP (15)	GCGACCCGAAGCACACCT
240	Gonyaulax PYK	GCCGAAGTTCACGATGTC
241	Euglena FBP (21)	CGGGCGACCTCCAGCATA
242	Gonyaulax PRK	CGACTACCACACCAACGA
244	Chondrus 6PGDH (15)	GCAGCCGCCCTTCCAAAT
245	Chondrus G6PDH (9)	ACTCTCTTGCCGCTGCTG
246	Chondrus G6PDH (9)	GCCCTCATCATCGGAAAT
249	Pavlova PRK	CTCCTGCATGTCGCGCTGG
250	Pavlova FBA (06-F02)	GGTGGCGAGCTTGTTTCATC
251	Chondrus TPI	ATGCGTAAGTGAATGTCC
252	Chondrus 6PGDH (15)	ATCAATCGCAACAAATCG
253	Phaeodactylum GPI	GACCAATCAACGCCAAAG
254	Phaeodactylum GPI	GTCCCGCACAAGGTTTTC
255	Gonyaulax PYK	TTCAAAGGCGACAACACC
256	Gonyaulax PRK	TCCCGCCTATTTTCGTTGA
262	Pavlova FBA (06-F02)	GCGCTTGCCCAGGTCGTC
263	Pavlova ENO	AGCCAGCCGTGCCCATCC
264	Chondrus 6PGDH (8)	TGAGGTCCCAGTCGTTCT
265	Chondrus 6PGDH (8)	AGAACGACTGGGACCTCA
266	Porphyra GPI	CGCTTCTCCGTGTGTTCC
267	Pyrocystis PYK	TGCCAGACCACGGAGGA
268	Gonyaulax PYK	TGCCGGACCGTGGACTCG
269	Guillardia FBP	CGATGGTGAAGACAAACA
270	Guillardia FBP	ATCCCCCTCGACGGCTCTT
271	Gonyaulax PRK	GGTGCGGGTCTTCCAAAT
272	Gonyaulax PRK	ACATCGCGCTGGACCTTC
273	Prymnesium PRK	GTGAAGTTTGCGTGGAAG
274	Pavlova ENO	AGCAACGTGAGCGGATAC
275	Porphyra TPI	CATGCGCTTGCCGAGACG
291	Pavlova GAPDH (CLAP+)	TGCACSACSAACKGSGTSGCNCC
292	Pavlova GAPDH (CLAP-)	GGSGCSAGSCMGTTSGTSGTGCA
293	Pavlova GAPDH (WYDNE-)	GAGTASCCCCACTCGTTGTCRTACCA
294	Pavlova GAPDH (GINGF+)	GGCATCAACGGCTTCGGSCGSATYGG
295	Pavlova SBP	GTSGCSTTCGACCCSCTSGACGG
301	Pavlova ENO	GGTTGAAGCACGGCATCG
302	Prymnesium PRK	GGTTTCCGTCCTTGAGAT
302b	Chondrus PRK	GAACCTGATGCCTGGGTA
303	Chondrus PRK	ACCACATCATCGTCATCG
304	Prymnesium PRK	AGGCTGACCATAGACATA
325	Guillardia PGK (36)	CCCCTCCTCATCATACAA
326	Guillardia PGK (37)	AAGATGCCGTGCGTCAGC
327	Guillardia PGK (37)	CCACCGCAGAGGATGATG
331	Guillardia PGK (36)	CCCCACCACCCCGTTCT
335	Prymnesium PRK	TTGCCAGTGGGGTCATCG
336	Prymnesium PRK	CCGCCGCCATCTTGTCCTC
337	Guillardia PRK(-)	GAGGTGGGTCTTGTCGTT
338	Guillardia PRK(+)	CCCCCATCATCATCTTCG
339	Pavlova PRK	GTTGCGGTGTCGTCTTGG
340	Pavlova PRK	CGCGTCGGCCTACTCAGC

7. Material und Methoden

341	Pavlova PRK	GGGCGTCAAGGCGTTCGT
342	Bigelowiella PRK	CATGCTGCCTGAGACTGC
343	Bigelowiella PRK	AGGGTCGTTTCAGCCAAAG
344	Bigelowiella PRK	TGCGACTGAGACTACCTT
345	Bigelowiella PRK	GCCTAGAAGCAGAGTAAA
346	Bigelowiella PRK	CGATGTAGAAGGTGAAGT
347	Bigelowiella PRK	TGAAACTATCCACCCTAA
348	Chondrus PRK	GACAAGTCTTCGCACCAC
349	Chondrus PRK	TTCGCCCCGCAACCAACA
352	Pavlova PRK	ACCGTGCTGATGGCGACC
354	Guillardia PRK	CCATGCCTTCGCTCTCAG
355	Guillardia PRK	ACACAAGGAAAGGGACAG
356	Guillardia PRK	CTCTCGTCTGCTGCTCAC
357	Guillardia PRK	CCGCTGATCTTACTTTTC
358	Euglena PRK	CGATCAGCTCATTACGA
359	Euglena PRK	CCAGTTGGGTAAAGTCAG
360	Euglena PRK	GGCAATCCAGTCCCTCAT
361	Euglena PRK	AAAAGTGGATAGATGAGC
362	Gonyaulax PRK	GGCACGAGGCAAGAGTTT
363	Gonyaulax PRK	TCCCGTTCGGTCGTGCTG
364	Gonyaulax PRK	GGGGAAGTGACGGTGGTA
367	Phaeodactylum PRK	CTCGTCCTCACTTTTCATT
368	Phaeodactylum PRK	AACAAGGCATTTTCTACT
369	Phaeodactylum PRK	TTTGCTGTCTTTGCCTCT
384	FBA-I (+/GC)	CCSGGSATCAAGGTSGACAAGGG
387	Phaeodactylum FBAI(3)	TATCCCCGGCATCAAGGT
388	Phaeodactylum FBAI(3)	AGCAGTGAGCCTTCCAAC
391	Guillardia PRK	GTAGACCTCACGGACCTTGAG
416	Prymnesium SBP(29)+	CAACAAC TGGGCGGTGGG
417	Prymnesium SBP(29)-	TTTGCCATCCAGTAGTCG
420	Prymnesium SBP(29)+	CGTCCCCGACATCTGGTA
421	Prymnesium SBP(29)-	AGCGGGTCCCAGCAGATG
425	Prymnesium SBP(29)+	GAGAGCGTGTTTCGGAGAC
426	Prymnesium SBP(29)+	GTGCTGGCGGACGACCTG
427	Prymnesium SBP(29)-	GCACCGCGAGCCCCCTAT
454	Coleochaete GapB	CGAAGACGTCAACACTGC
455	Coleochaete GapB	GGCATTCTGGGGGTCTCA
456	Marchantia GapB	GAGAAGAAGGGCCTGACG
457	Marchantia GapB	ACCGTCGACTCCTCCCTC
458	Mesostigma GapB	CAATCCGCGACCATCAGC
459	Mesostigma GapB	TGCTATTTAGTCGTTGTA
460	Mesostigma GapB	GCGCGGATGAGGCTTTTCG
461	Mesostigma GapB	GCGGCACTCATCAGCATC
462	Mesostigma GapA1	CTGACAGCCTTGACATCG
463	Mesostigma GapA1	CGTGCCGAGGGTGAAATC
464	Mesostigma GapA2	CCGAGGCATTGGAGACGA
465	Mesostigma GapA2	GGGCGGTGGGAAAAACAT
466	Chara GapA	TCGCCCTTGAGATTGTG
467	Chara GapA	GCGTTGTTGACCTCTTCT
468	Marchantia GapB	ATCACCGTGGGAGCATCG
469	Marchantia GapB	GTCTCCGCTGGGCTTGGC
470	Marchantia GapB	GCAGCTCGCTCGTCCAAT
471	Marchantia GapB	GCCATTGCGCTGCTACCA

7. Material und Methoden

472	Marchantia GapA	ACTATGTCCGTCGCCTCG
473	Marchantia GapA	CGAGACCTACGGGGATGT
474	Marchantia GapA	TCCGACGACCTGTTGACC
475	Marchantia GapA	GCATGGTCGGTGTTTTAT
476	Prymnesium SBP(29)	TCCTCCTCACTCGCTCCTTCCTT
477	Prymnesium SBP(29)	ACTTCCCACATCAGGTCGTC
478	Pavlova PGK	GACACCTTTGAGCCGCCGACGAT
479	Pavlova PGK	GGCTTGAGGCACCTTGGTC
480	Prymnesium PGK	GGGGCGCTTGGGGTGGACAC
481	Prymnesium PGK	GCCATCCAAGTAGTCCAG
482	Prymnesium TKL	TGGTTGTCGTCGTAGAAGG
483	Prymnesium TKL	CCATGATGCAGTAGGTGT
503	Cladophora GapA	TTAGACACATATAGATCA
504	Cladophora GapA	ACAGCCACGCAAACAGTA
505	Cladophora GapA	GCAGACTTCTATGGTTGG
506	Cladophora GapA	GCCATCACAGCATCATTC
507	Mesostigma GapB	GCAGCCAGAAGCAGGTGA
508	Mesostigma GapB	GTCGTGGGTGTAGTCCTT
509	Mesostigma GapB	CTTGGCGGGGCGGTGAT
510	Mesostigma GapB	CCGGCAAGCACATCATCG
511	Mesostigma GapB	GGTGCCAAGAAGGTGCTC
512	Mesostigma GapB	ATGCCGGGCCAGTTGTTG
513	Marchantia GapA	AGAGTGGGTGGTGGTCAT
514	Marchantia GapB	CTTCAGCAGTTACCTTCG
515	Marchantia GapB	TCGGAGCAGTGGAAGTCG
521	Marchantia GapB	CATCAAGGGCGAGACCAG
522	Chara GapB	CCTCGTGCTCCCCAAGTT
523	Marchantia CP12	AATTTGTTTCGTCGCATCG
524	Marchantia CP12	GGGTTGCTGTCAATTCCTC
525	Marchantia GapC	CACCTTCCCATCATTCCTT
526	Marchantia GapC	TTGGTTCTCACTGGGTAG
527	Marchantia GapC	TGTTTTTGACGGTTTTGTC
528	Marchantia GapC	ATGCAAGATCAGGTCCAC
529	Marchantia GapCp	ACACTTCCACTATTTTCG
530	Marchantia GapCp	TAGTTCTTCTTAGCGGAG
531	Marchantia GapCp	GATTTCTTCTACCCTTGC
532	Marchantia GapCp	AAC TGCTGCGATGTGCTC
533	Coleochaete GapC	AGGGCGTGATGATGTTGA
534	Coleochaete GapC	AAACAAGCTTCACAAACT
535	Coleochaete GapC	CTTCATCAGCACGGAGTA
536	Coleochaete GapC	ATTCCAGCTTTCGCATCG
545	Marchantia GapC	CGAGTTGGTTGCCGTGAA
546	Marchantia GapCp	TGATGAGTCTTCGGTCTT
547	Gonyaulax FBP	GTTCGACCCGTTGGATGG
548	Gonyaulax FBP	GATGTAGCGGAGGGAGTA
549	Marchantia GapC	GATAGCGGCCTTGACTTC
550	Marchantia GapCp	ACATGCAAACCAAGGGAT
554	Coleochaete GapA	GGGTTTGGTCGCATTGGA
555	Coleochaete GapA	ATGGAAGGAAGGACTCTC
556	Coleochaete GapA	TACCATGCAACCACCTTA
557	Coleochaete GapA	TGTCATCACCATAACCA
558	Coleochaete GapB	CGTATTGGTCGTAACTTT
559	Coleochaete GapB	TGGTATGAGCGGGAAGAC

7. Material und Methoden

560	Coleochaete GapB	ACCCCATTCGTTGTCGTA
561	Coleochaete GapB	GTCATCACCCATCACCAT
562	Gonyaulax FBACyt	ATGTGCGAGGAGAAGAGC
563	Coleochaete GapC	TGGGGTTTACACAGACAT
564	Coleochaete GapC	TGGTAACTTTGGGGAACG
579	Coleochaete GapC	TCAGCACTCAGATTCATA
580	Chara GapB	CGTGGACGGGAAAGTGAT
581	Chlorokybus GapA	GGTGCTGGTGGGGTTCGTG
582	Chlorokybus GapA	CCCGATGTCAAGGCTGTT
583	Chlorokybus GapA	CGACGGTGGAGTTCAGTT
584	Chlorokybus GapA	GGCGACGATGTGGCTGAG
585	Chlorokybus GapB	ATGGCGGCGCAACACTCA
586	Chlorokybus GapB	GCGCCACCCCCACCAAGG
587	Chlorokybus GapB	CGTCTGCCTCGGGTTCG
588	Chlorokybus GapB	CGCCACTCTCCTCCTCAC
589	Marchantia SBP	CCGAGGAGGAACCCAACC
590	Marchantia SBP	GTGTACTTCTCGCTGACG
591	Marchantia FBP	AGGTAAC TACGCACTCTG
592	Marchantia FBP	GTGGCACTCGCTGATGTA
595	Marchantia FBACyt	AACGGAGAAACAACAAC
596	Marchantia FBACyt	TAACCGAGCCGTCTGAAC
597	Marchantia FBACipla	GATGGAACCGATGGAGAA
598	Marchantia FBACipla	AAACACCCTTCCTCCACA
601	Marchantia FBACyt	GTGTTGTTTCGTTTAGAGC
602	Marchantia FBACipla	GGCATGAGAGTTGACCTT
603	Marchantia CP12	GGCGTAAGCGGCGGACAG
605	Coleochaete GapC	GAGAGGCGATTGTGTCCA
606	Chara GapB	TGGTTATGGGCGATGACA
607	Coleochaete GapC	GCCGTGGTGCTGGTTTTTA
608	Marchantia FBACyt	TGAAACTAACAGGCAAGC
609	Marchantia FBACyt	TGAGAACCATCCACTAAT
649	Marchantia GPI	TTCTGGGTTTGATGAGTG
653	Chara GapA	ATTGGTCGTAACCTCCTG
654	Chara GapA	ATGCTGGCACACAAGAAA
655	Chara GapA	ATCGTCACCCATAACCAT
656	Chara GapB	TTCGCCATCTCCATAGCC
657	Chara GapB	TCCTCCCCCTCGTTCACCT
658	Chara GapB	ACTCCGAATGGCACTGAA
659	Chara GapB	TCGAAAACACGGCACTCC
660	Marchantia CP12	CGTCCATCTTGCGTCTAA
661	Marchantia CP12	GCGGAGGGAGTTGGATTG
667	Marchantia SBP	GGGCCACCGACTTCACTG
668	Marchantia SBP	GACGGAACCTGGCAACAC
672	Marchantia GPI	GTCGTCTGTGCTTCTTCCA
673	Marchantia GPI	CTCATCAAACCCAGAAGA
674	Marchantia GPI	GTTTCATCCTCGTCTCTCA
675	Marchantia GPI	GAGCTGAGCGGAAATGAT
676	Marchantia GPI	TTTGGATTCTGGGACTGG
677	Marchantia GPI	AGCCAGGGACTCATCAGA
678	Marchantia GPI	GTTGGTGGACGATACAGT
679	Marchantia GPI	ATGGGGGTCTGGATACAA
680	Marchantia SBP	CAGACAGCAAGGAGGAGA
681	Marchantia SBP	CGTGGCTGCATTTGAACT

7. Material und Methoden

682	Marchantia SBP	AGGGAAGGGAAAGGAATC
683	Marchantia SBP	CTCTGCCGAAACTACCTA
684	Marchantia FBAI	AACCCCTTTCGCAGATGAG
685	Marchantia FBAI	CTTGCTGTTCCTCTTGC
686	Marchantia FBAI	AGCGCCCGGGAAAGGAAT
687	Marchantia FBAI	GAGTAGCCCGCATTAGTA
689	Chondrus FBPCyt	TGAAGGAACAGGAAGACG
690	Chondrus FBPCyt	AGCTTCAAAGTCTGCAAT
691	Chondrus FBPCyt	GGACAGCGCAAGCACCAC
692	Chondrus FBPCyt	ATCTTCCGACGACCCCAT
695	Guillardia SBP	ATGGCGTTTCCTTCCTGTG
696	Guillardia SBP	GCTCCTTATCGGCGTCAT
697	Guillardia SBP	TTTTACTTGCCGAGAACC
698	Guillardia SBP	CAGGAGTTGCCAGTGACG
699	Chondrus SBP	CAGTCTCCCTCATCAACA
700	Chondrus SBP	TCGCAATCGCACCCACTA
701	Chondrus SBP	TCGGAAAACAAGACAAGT
702	Chondrus SBP	CGTGACACCAACCAACTC
703	Chara GapB	ATGTGTTTCGGGGGTGTCA
704	Chara GapB	GCTAAGGGTGGCAATCAA
705	Gonyaulax FBPpla	ACTGGGCTGTCTCCTCC
706	Gonyaulax FBPpla	ACGCCTTCTTGAGCACAT
707	Gonyaulax SBP	CCAGGGTGAACCTCGTAGA
708	Gonyaulax SBP	GTCCCAAACGCCCACAAT
713	Marchantia SBP	CACAAAACACGGACAGAC
714	Marchantia SBP	GATGGTGGTGGTTTCCTT
715	Marchantia FBA	CGTGACAGCGGGAATACT
720	Marchantia FBA	GCCTGACGACGAGAGTAG
721	Chlorokybus GapB	GCACCTTCAACGCCAACG
724	Marchantia FBA	CAGAGACTCGTGGCATT
737	Chondrus SBP	CTTTGACCCCTTGGATGG
738	Chara GapA	CAACAGAAACCCGCTCAA
739	Mesostigma FBAI	GGCGGCTACCAACAAGAG
740	Mesostigma FBAI	ACTGGTATCCCTTCTCGT
741	Mesostigma FBAI	GCTTGTACTIONGCCCATCT
742	Mesostigma SBP	TCGTCTTTGCGGCTACCT
743	Mesostigma SBP	ACCGTTTAATCTCCCTAT
744	Mesostigma SBP	CATGCAGGCTTCCCAACG
745	Mesostigma SBP	CAGCACGATGAAGAGGTC
746	Mesostigma GapC	ATTAAATCCGCATGGCT
747	Mesostigma GapC	ATCGTCGCCGTCTACTA
748	Mesostigma GapC	AAGATCGGCATCAATGGT
749	Mesostigma GapC	TCAACAGGGACCACAATG
750	Mesostigma GapA	TTGGCCGCAACTTCCTG
751	Mesostigma GapA	CGTTTCAAAATGTCAC TG
752	Mesostigma GapA	TGGCACGGTCGGGAGAAC
753	Mesostigma GapA	CGGCACAATTTACTTCAT
754	Mesostigma CP12	GATATGGCTACCGCAACC
755	Mesostigma CP12	TGCTTTTAGACCCTTACG
756	Mesostigma CP12	CGCAGTTAGTTTGAGGTC
757	Mesostigma CP12	ATGCGGGTGGAATGTAGC
758	Chlorokybus GapA	GGCTTCGGACGCATTGGC
759	Chlorokybus GapA	AAGTTATCGCAGTTGGTC

7. Material und Methoden

760	Chlorokybus GapA	TGCTGGTGGGGTCGTGAG
761	Chlorokybus GapA	GCTCAAGCTGGGACTCTA
764	Chara GapA	CGAGGGGTCATAATCTGC
778	Gonyaulax SBP (Mini5)	GCCCAGGGTGAACCTCGTA
779	Mesostigma SBP	TCAAGTGGCAGTCTCTCG
780	Mesostigma SBP	GAGCCGTCCAGGGGGTCA
781	Mesostigma GapB	CTCCCTCTTGCCGTTTTTC
782	Mesostigma GapC	TGGTCGGCGGTGCTGAGA
783	Mesostigma GapC	GTCATCTCCGCCCCATCA
798	Mesostigma SBP (genomisch)	CTGCAATTTGTAGACCTG
799	Mesostigma GapB (genomisch)	AAGGAGGAGGAGCAGCAG
800	Prymnesium SBP	TCCTGCTTCTCCGCCACC
817	Mesostigma GapC	GGTGCTGCCTGCCCTGAA
818	Mesostigma GapC	CGTGCCACCCCCGATGT
819	Mesostigma GapC	ACGACGCACCCTTCTCCA
820	Mesostigma GapC	ACGGCATCCTCGGTGTAG
821	Mesostigma GapC	TGCTTGTTGCCGATGAA
828	Prymnesium PRK	CGTCCTCCCACCCTCCAG
829	Prymnesium PRK	ATGCTCTCCCTCTCTGTC
830	Prymnesium PRK	TATTTTCGTGCCAGTATTG
831	Prymnesium PRK	AGCCTTGTTGTTTACCAT
832	Prymnesium SBP	CTCCTCCCCTCCCCTCGTG
833	Prymnesium SBP	CAGCTTCTACACCGATGC
834	Prymnesium SBP	TGAGCATCGTGTTGAAGC
835	Prymnesium SBP	ACACGCCATCGGAGGTCT
828	Prymnesium PRK	CGTCCTCCCACCCTCCAG
829	Prymnesium PRK	ATGCTCTCCCTCTCTGTC
830	Prymnesium PRK	TATTTTCGTGCCAGTATTG
831	Prymnesium PRK	AGCCTTGTTGTTTACCAT
832	Prymnesium SBP	CTCCTCCCCTCCCCTCGTG
833	Prymnesium SBP	CAGCTTCTACACCGATGC
834	Prymnesium SBP	TGAGCATCGTGTTGAAGC
835	Prymnesium SBP	ACACGCCATCGGAGGTCT
856	Mesostigma GapC	GGTAGGCTGCTGAGGAGT
869	Marchantia GapB	CGGTGTAAGAGTGGGTGG
880	Mesostigma GapC	CGCCGAGGGGATGATGTT
937	Mesostigma GapC	CACGCTCACCGCCTGCTC
939	Coleochaete GapC	CAGGGAGAAAGAACAGAT

8. Literaturverzeichnis

- Archibald J.M.** und Keeling P.J. (2002). Recycled plastids: a 'green movement' in eukaryotic evolution. *Trends in Genetics* 18 (11): 577-584.
- Archibald J.M.**, Rogers M.B., Toop M., Ishida K. und Keeling P.J. (2003). Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigeloviella natans*. *Proc. Natl. Acad. Sci. USA* 100 (13): 7678-7683
- Baptiste E.** und Philippe H. (2002). The Potential Value of Indels as Phylogenetic Markers: Position of Trichomonads as a Case Study. *Mol. Biol. Evol.* 19(6):972–977.
- Berget S.M.**, Moore C. und Sharp P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* 74 (8): 3171-3175.
- Baldauf S.L.**, Roger A.J., Wenk-Siefert I. und Doolittle W.F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290 (5493): 972-977.
- Bhattacharya D.** und Medlin L. (1995). The phylogeny of plastids: a review based on comparison of small-subunit ribosomal RNA coding regions. *J. Phycol.* 31:489-498.
- Bhattacharya D.**, Weber K., An S.S. und Berning-Koch W. (1998). Actin phylogeny identifies *Mesostigma viride* as a flagellate ancestor of the land plants. *J. Mol. Evol.* 47: 544-550.
- Bhattacharya D.** (1998). Algal phylogeny and the origin of land plants. *Plant Physiol.* 116: 9-15.
- Biesecker G.**, Harris J.I., Thierry J.C., Walker J.E. und Wonacott A.J. (1977). Sequence and structure of D-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Nature* 266: 328-333.
- Bremer K.** (1985). Summary of green plant phylogeny and classification. *Cladistics* 1: 369-385.
- Bremer K.**, Humphries C. J., Mishler D. und Churchill S. P. (1987). On cladistic relationships in green plants. *Taxon* 36: 339–349.
- Brenniker A.**, Grohmann L., Hiesel R., Knoop V. und Schuster W. (1993). The mitochondrial genome on its way to the nucleus: different stages of gene transfer in higher plants. *FEBS Lett.* 325 (1-2): 140-145.
- Brinkmann H.**, Cerff R., Salomon M. und Soll J. (1989). Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. *Plant Mol. Biol.* 13: 81-94.
- Brinkmann H.**, van der Giezen M., Zhou Y., Poncelin de Raucourt G. und Philippe H. (2005). An empirical assessment of long-branch attraction artefacts in deep eukaryotic phylogenomics. *Syst. Biol.* 54: 743-757.

8. Literaturverzeichnis

- Buchanan B.B.** (1980). Role of light in the regulation of chloroplast enzymes. *Annu Rev Plant Physiol.* 31:341-374.
- Cavalier-Smith T.** (1981). Eukaryote kingdoms: seven or nine? *Biosystems* 14: 461–481.
- Cavalier-Smith T.** (1985). Selfish DNA and the origin of introns. *Nature* 315: 283–284.
- Cavalier-Smith T.** (1999). Principles of Protein and Lipid Targeting in Secondary Symbiogenesis: Euglenoid, Dinoflagellate, and Sporozoan Plastid Origins and the Eukaryote Family Tree. *Journal of Eukaryotic Microbiology* 46 (4): 347–366.
- Cerff R.** (1979). Quaternary structure of higher plant glyceraldehyde-3-phosphate dehydrogenases. *Eur. J. Biochem.* 94: 243-247.
- Cerff R.** und Chambers S.E. (1979). Subunit structure of higher plant glyceraldehyde-3-phosphate dehydrogenases. *J. Biol. Chem.* 254: 6094-6098.
- Cerff R.** (1982). Separation and purification of NAD- and NADP-linked glyceraldehyde-3-phosphate dehydrogenases from higher plants. In Edelman M, Hallick RB, Chua N-H (eds.) *Methods in Chloroplast Molecular Biology*, pp. 683-694. Elsevier Biomedical Press, Amsterdam.
- Cerff R.** (1995). “The chimeric nature of nuclear genomes and the antiquity of introns as demonstrated by the GAPDH gene system.” in M. Go and P. Schimmel, eds. *Tracing biological evolution in protein and gene structures*. Elsevier, Amsterdam. Pp. 205–227.
- Chesnick J.M.,** Kooistra W.H., Wellbrock U. und Medlin L.K. (1997). Ribosomal RNA analysis indicates a benthic pennate diatom ancestry for the endosymbionts of the dinoflagellates *Peridinium foliaceum* and *Peridinium balticum* (Pyrrhophyta). *J Eukaryotic Microbiol* 44 (4): 314-320
- Chow L.T.,** Gelinas R.E., Broker T.R. und Roberts R.J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12 (1): 1-8.
- Ciniglia C.,** Yoon H.S., Pollio A., Pinto G. und Bhattacharya D. (2004). Hidden biodiversity of the extremophilic Cyanidiales red algae. *Molecular Ecology* 13: 1827–1838.
- Clermont S.,** Corbier C., Mely Y., Gerard D., Wonacott A. und Branlant G. (1993). Determinants of Coenzyme Specificity in Glyceraldehyde-3-phosphate Dehydrogenase: Role of the Acidic Residue in the Fingerprint Region of the Nucleotide Binding Fold. *Biochemistry* 32: 10178-10184.
- Coppin A.,** Varre J.S., Lienard L., Dauvillee D., Guerardel Y., Soyer-Gobillard M.O., Buleon A., Ball S. und Tomavo S. (2005). Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. *J. Mol. Evol.* 60: 257-267.
- Daugbjerg N.** und Andersen R.A. (1997). Phylogenetic analysis of the *rbcL* sequences from haptophytes an the heterokont algae suggest their chloroplasts are unrelated. *Mol. Biol. Evol.* 14: 1242-1251.

8. Literaturverzeichnis

- Delwiche C.F.** (1999). Tracing the Thread of Plastid Diversity through the Tapestry of Life. *The American Naturalist* 154: 164–177.
- de Souza S.J.**, Long M., Klein R.J., Roy S., Lin S. und Gilbert W. (1998). Toward a resolution of the introns early/late debate: Only phase zero introns are correlated with the structure of ancient proteins (introns-early/introns-late/modules). *Proc. Natl. Acad. Sci. USA* 95: 5094–5099.
- Douglas S.E.** (1998). Plastid evolution: origins, diversity, trends. *Current Opinion in Genetics & Development* 8: 655–661.
- Douglas S.**, Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T., Wu X., Reith M., Cavalier-Smith T. und Maier U.G. (2001). The highly reduced genome of an enslaved algal nucleus. *Nature* 410: 1091–1096.
- Douzery E.J.P.**, Snell E.A., Baptiste E., Delsuc F. und Philippe H. (2004). The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. USA* 101 (43): 15386–15391.
- Durnford D.G.**, Deane J.A., Tan S., McFadden G.I., Gantt E. und Green B.R. (1999). A phylogenetic assessment of the eucaryotic light-harvesting antenna proteins, with implications for plastid evolution. *J. Mol. Evol.* 48: 59–68.
- Dzierszinski F.**, Popescu O., Toursel C., Slomianny C., Yahiaoui B. und Tomavo S. (1999). The Protozoan Parasite *Toxoplasma gondii* Expresses Two Functional Plant-like Glycolytic Enzymes. *The Journal of Biological Chemistry* 274 (35): 24888–24895.
- Edwardsen B.**, Eikrem W., Green J.C., Andersen R.A., Moon-van der Staay S.Y. und Medlin L.K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia* 39(1): 19–35.
- Fagan T.**, Hastings J.W. und Morse D. (1998). The phylogeny of glyceraldehyde-3-phosphate dehydrogenase indicates lateral gene transfer from cryptomonads to dinoflagellates. *J. Mol. Evol.* 47: 633–639.
- Falkowski P.G.**, Katz M.E., Knoll A.H., Quigg A., Raven J.A., Schofield O. und Taylor F.J.R. (2004). The Evolution of Modern Eukaryotic Phytoplankton. *Science* 305: 354–360
- Fast N.M.**, Kissinger J.C., Roos D.S., Keeling P.J. (2001). Nuclear-encoded, plastid targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol. Biol. Evol.* 18: 418–426.
- Fawley M.W.**, Yun Y. und Qin M. (2000). Phylogenetic analyses of 18S rDNA sequences reveal a new coccoid lineage of the Prasinophyceae (Chlorophyta). *J. Phycol.* 36: 387–393.
- Fedorov A.** und Fedorova L. (2004). Introns: Mighty Elements from the RNA World. *J. Mol. Evol.* 59:718–721.
- Field C.B.**, Behrenfeld M.J., Randerson J.T. und Falkowski P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281 (5374): 237–240.

8. Literaturverzeichnis

- Figge R.M.**, Schubert M., Brinkmann H. und Cerff R. (1999). Glyceraldehyde-3-phosphat dehydrogenase gene diversity in eubacteria and eucaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol. Biol. Evol.* 16: 429-440.
- Foth B.J.** und McFadden G.I. (2003). The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. *Int. Rev. Cytol.* 224: 57-110.
- Fothergill-Gilmore L.A.** und Michels P.A. (1993). Evolution of glycolysis. *Progress in Biophysics and Molecular Biology* 59 (2): 105-235.
- Frischauf A.M.**, Lehrach H., Poustka A. und Murray N. (1983). *Journal of Molecular Biology* 170: 827.
- Frohman M.A.** (1993). Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Methods Enzymol.* 218: 340-356.
- Funes S.**, Davidson E., Reyes-Prieto A., Magallon S., Herion P., King M. und Gonzalez-Halphen D. (2002). A green algal apicoplast ancestor. *Science* 298: 2155.
- Funes S.**, Davidson E., Reyes-Prieto A., Magallon S., Herion P., King M. und Gonzalez-Halphen D. (2003). Response to comment on "A green algal apicoplast ancestor". *Science* 301: 49b.
- Gerbling K.P.**, Steup M. und Latzko E. (1986). Fructose-1,6-bisphosphatase form B from *Synechococcus leopoliensis* hydrolyzes both fructose an sedoheptulose bisphosphate. *Plant Physiol.* 80: 716-720.
- Gibbs S.P.** (1978). The chloroplast of *Euglena* may have evolved from symbiotic green algae. *Can. J. Bot.* 56, 2883–2889.
- Guillou L.**, Eikrem W., Chretiennot-Dinet M.J., Le Gall F., Massana R., Romari K., Pedros-Alio C. und Vaultot D. (2004). Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* 155:193–214.
- Graham L.E.** und Wilcox L.W. (2000). *Algae*. Prentice-Hall Inc., New York.
- Graham L.E.**, Cook M.E. und Busse J.S. (2000). The origin of plants: Body plan changes contributing to a major evolutionary radiation. *Proc. Nat. Acad. Sci. USA* 97 (9): 4535–4540.
- Gray M.W.**, Lang B.F., Cedergren R., Golding G.B., Lemieux C., Sankoff D., Turmel M., Brossard N., Delage E., Littlejohn T.G., Plante I., Rioux P., Saint-Louis D., Zhu Y. und Burger G. (1998). Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Research* 26 (4): 865–878.
- Greener A.** (1990). *Strategies* 3: 5-6.
- Gross W.**, Bayer M.G., Schnarrenberger C., Gebhart U.B., Maier T.L. und Schenk H.E.A. (1994). Two Distinct Aldolases of Class II Type in the Cyanoplasts and in the Cytosol of the Alga *Cyanophora paradoxa*. *Plant Physiol.* 105: 1393-1398.

8. Literaturverzeichnis

- Gross W.**, Lenze D., Nowitzki U., Weiske J. und Schnarrenberger C. (1999). Characterization, cloning, and evolutionary history of the chloroplast and cytosolic class I aldolases of the red alga *Galdieria sulphuraria*. *Gene* 230:7–14.
- Hacket J.D.**, Maranda L., Yoon H.S. und Bhattacharya D. (2003). Phylogenetic Evidence for the Cryptophyte Origin of the Plastid of *Dinophysis* (Dinophysiales, Dinophyceae). *J. Phycol.* 39: 440–448.
- Hannaert V.**, Brinkmann H., Nowitzki U., Lee J.A., Albert M.-A., Sensen C.W., Gaasterland T., Müller M., Michels P. und Martin W. (2000). Enolase from *Trypanosoma brucei*, from the Amitochondriate Protist *Mastigamoeba balamuthi*, and from the Chloroplast and Cytosol of *Euglena gracilis*: Pieces in the Evolutionary Puzzle of the Eukaryotic Glycolytic Pathway. *Mol. Biol. Evol.* 17 (7): 989–1000.
- Hannaert V.**, Saavedra E., Duffieux F., Szikora J.P., Rigden D.J., Michels P.A.M. und Opperdoes F.R. (2003). Plant-like traits associated with metabolism of *Trypanosoma* parasites. *Proc. Natl. Acad. Sci. USA* 100 (3): 1067–1071.
- Harper J.T.** und Keeling P.J. (2003). Nucleus-Encoded, Plastid-Targeted Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) Indicates a Single Origin for Chromalveolate Plastids. *Mol. Biol. Evol.* 20(10): 1730–1735.
- Harrison D.H.T.**, Runquist J.A., Holub A. und Miziorko H.M. (1998). The Crystal Structure of Phosphoribulokinase from *Rhodobacter sphaeroides* Reveals a Fold Similar to That of Adenylate Kinase. *Biochemistry* 37: 5074–5085.
- Hartleben S.E.** (2005). Klasse II Aldolasen als Marker für einen monophyletischen Ursprung der Plastiden von Chromisten und Dinoflagellaten: Molekulare Untersuchung und phylogenetische Charakterisierung der FBAIL-Gene. Diplomarbeit.
- Häuber M.M.**, Müller S.B., Speth V. und Maier U.G. (1994). How to evolve a complex plastid? – A hypothesis. *Bot. Acta* 107: 383–386.
- Huse W.D.** und Hansen C. (1988). Strategies 1: 1–3.
- Junk S.V.** (2005). Lokalisierung von Fructose-1,6-bisphosphat Aldolasen der Klasse II aus komplexen Algen in der Diatomee *Phaeodactylum tricornutum* mit Hilfe von GFP-Reportergenkonstrukten. Diplomarbeit.
- Jurica M.S.** und Moore M.J. (2003). Pre-mRNA Splicing: Awash in a Sea of Proteins. *Molecular Cell* 12: 5–14.
- Karol K.G.**, McCourt R.M., Cimino M.T. und Delwiche C.F. (2001). The Closest Living Relatives of Land Plants. *Science* 294: 2351–2353.
- Keeling P.J.** und Doolittle W.F. (1997). Evidence that eukaryotic triosephosphate isomerase is of alpha-proteobacterial origin. *Proc. Natl. Acad. Sci. USA* 94: 1270–1275.
- Keeling P.J.** und Palmer J.D. (2001). Lateral transfer at the gene and subgenic levels in the evolution of eukaryotic enolase. *Proc. Natl. Acad. Sci. USA* 98 (19): 10745–10750.

8. Literaturverzeichnis

- Keeling P.J.**, Archibald J.M., Fast N.M. und Palmer J.D. (2004). Comment on: "The Evolution of Modern Eukaryotic Phytoplankton". *Science* 306: 2191.
- Kersanach R.**, Brinkmann H., Liaud M.F., Zhang D.X., Martin W. und Cerff R. (1994). Five identical intron positions in ancient duplicate genes of eubacterial origin. *Nature* 367: 387-389.
- Kilian O.** und Kroth P.G. (2004). Presequence Acquisition During Secondary Endocytobiosis and the Possible Role of Introns. *J. Mol. Evol.* 58: 712–721.
- Kilian O.** und Kroth P.G. (2005). Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *The Plant Journal* 41: 175–183.
- Kim Y.J.**, Lee S.H. und Park K.Y. (2004). A leader intron and 115-bp promoter region necessary for expression of the carnation S-adenosylmethionine decarboxylase gene in the pollen of transgenic tobacco. *FEBS Letters* 578 (2004) 229–235
- Köhler U.**, Liaud M.F., Mendel R.R., Cerff R. und Hehl R. (1995). The maize GapC4 promoter confers anaerobic reporter gene expression and shows homology to the maize anthocyanin regulatory locus C1. *Plant Mol. Biol.* 29 (6): 1293-1298.
- Kroth P.G.**, Schroers Y. und Kilian O. (2005). The peculiar distribution of class I and class II aldolases in diatoms and in red algae. *Curr. Genet.* 48: 389–400.
- Lefebvre S.**, Lawson T., Zakhleniuk O.V., Lloyd J.C., Raines C.A. und Fryer M. (2005). Increased sedoheptulose-1,7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. *Plant Physiol.* 138: 451-460.
- Leister D.** und Schneider A. (2004). Evolutionary Contribution of Plastid Genes to Plant Nuclear Genomes and Its Effects on Composition of the Proteomes of All Cellular Compartments. [Beitrag im Sammelband] *Organelles, Genomes and Eukaryote Phylogeny: An Evolutionary Synthesis in the Age of Genomics* 68: 233-251.
- Lemieux C.**, Otis C. und Turmel M. (2000). Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature* 403: 649–652.
- Liaud M.F.**, Brandt U., Scherzinger M. und Cerff R. (1997). Evolutionary Origin of Cryptomonad Microalgae: Two Novel Chloroplast/ Cytosol-Specific GAPDH Genes as Potential Markers of Ancestral Endosymbiont and Host Cell Components. *J.Mol. Evol.* 44: 28-37.
- Liaud M.F.**, Lichtlé Ch., Apt K., Martin W. und Cerff R. (2000). Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eucaryotic glycolytic pathway. *Mol. Biol. Evol.* 17: 213-223.
- Löffelhardt W.**, Bohnert H.J. und Brynatt D.A. (1997). The complete sequence of the *Cyanophora paradoxa* cyanobacterial genome (Glaucocystophyceae). *Plant Syst. Evol.* 11(Suppl.): 149–162.

8. Literaturverzeichnis

- Long M.**, Cerff R. (2003). Introns: Movements. in Nature Encyclopedia of the human genome. Macmillan Publishers Ltd, Nature Publishing Group. Pp. 547-551.
- Long M.**, de Souza SJ, Rodenberg C., Gilbert W. (1996). Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor. Proc. Natl. Acad. Sci. USA 93: 7727-7731.
- Lüttge U.**, Kluge M., Bauer G. (1999). Botanik (3. Auflage). Wiley-VCH Verlag GmbH.
- Lynch M.**, Conery J. S. (2000). The evolutionary fate and consequences of duplicate genes. Science 290: 1151–1155.
- Malin G.** und Kirst G.O. (1997). Algal production of dimethyl sulfide and its atmospheric role. Journal of Phycology 33: 889-896.
- Margulis L.**, Schwartz K.V. (1989). Die fünf Reiche der Organismen. Spektrum Verlag
- Marin B.**, Melkonian M. (1999). Mesostigmatophyceae, a new class of streptophyte green algae revealed by SSU rRNA sequence comparisons. Protist 150: 399–417.
- Martin W.**, Brinkmann H., Savona C., Cerff R. (1993). Evidence for a chimaeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenases. Proc. Natl. Acad. Sci. USA 90: 8692-8696.
- Martin W.**, Mustafa A.Z., Henze K., Schnarrenberger C. (1996). Higher-plant chloroplast and cytosolic fructose-1,6-bisphosphatase isoenzymes: origins via duplication rather than prokaryote-eukaryote divergence. Plant Mol. Biol. 32: 485-491.
- Martin W.**, Schnarrenberger C. (1997). The evolution of the Calvin Cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr. Genet. 32: 1-18.
- Martin W.**, Herrmann R.G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. 118: 9-17
- Martin W.**, Stöbe B., Goremykin V., Hansmann S., Hasegawa M., Kowallik K.V. (1998). Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393: 162-165.
- Martin W.**, Rujan T., Richly E., Hansen A., Cornelsen S., Lins T., Leister D., Stoebe B., Hasegawa M., Penny D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA 99: 12246–12251.
- Matsuzaki M.**, Misumi O., Shin-i T., Maruyama S., Takahara M., Miyagishima S., Mori T., Nishida K., Yagisawa F., Nishida K., Yoshida Y., Nishimura Y., Nakao S., Kobayashi T., Momoyama Y., Higashiyama T., Minoda A., Sano M., Nomoto H., Oishi K., Hayashi H., Ohta F., Nishizaka S., Haga S., Miura S., Morishita T., Kabeya Y., Terasawa K., Suzuki Y., Ishii Y., Asakawa S., Takano H., Ohta N., Kuroiwa H., Tanaka K., Shimizu N., Sugano S., Sato N., Nozaki H., Ogasawara N., Kohara Y. und Kuroiwa T. (2004). Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428: 653-657.

8. Literaturverzeichnis

- Mattox K.R.**, Stewart K. D. (1984). Classification of the green algae: a concept based on comparative cytology. Pp. 29–72 in D.E.G. Irvine and D.M. John, eds. The systematics of green algae. Academic Press, London.
- McCourt R.M.**, Delwiche C.F., Karol K.G. (2004). Charophyte algae and land plant origins. Trends Ecol. Evol. 19: 661–666.
- McFadden G.I.**, Gilson P.R. und Waller R.F. (1995). Molecular phylogeny of Chlorarachniophytes based on plastid rRNA and rbcL sequences. Arch. Protistenkd. 145: 231–239.
- McFadden G.I.**, Gilson P.R. und Hofman C.J. (1997). Division Chlorarachniophyta. in: Bhattacharya D. (ed) Origins of algae and their plastids. Springer Verlag, Wien, pp 175–185.
- Melkonian M.** (1989). Flagellar apparatus ultrastructure in *Mesostigma viride* (Prasinophyceae). Plant Syst. Evol. 164: 93–122.
- Mereschkowsky C.** (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. Biologisches Centralblatt No. 10: 321–347.
- Meyer-Gauen G.**, Schnarrenberger C., Cerff R. und Martin W. (1994). Molecular characterization of a novel, nuclear-encoded, NAD(1)-dependent glyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L. Plant Mol. Biol. 26: 1155–1166.
- Michels A.**, Wedel N. und Kroth P.G. (2005). Diatom Plastids Possess a Phosphoribulokinase with an Altered Regulation and No Oxidative Pentose Phosphate Pathway. Plant Physiology 137: 911–920.
- Miyagawa Y.**, Tamoi M., Shigeoka S. (2001). Overexpression of a cyanobacterial fructose-1,6- /sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. Nat. Biotechnol. 19: 965–969.
- Morden C.W.** und Sherwood A.R. (2002). Continued evolutionary surprises among dinoflagellates. Proc. Natl. Acad. Sci. USA 99 (18): 11558–11560.
- Morse D.**, Salois P., Markovic P., Hastings J.W. (1995). A nuclear-encoded formII RuBisCO in dinoflagellates. Science 268: 1622–1624.
- Murray N.E.** (1983). Lambda II. In Lambda II (New York: Cold Spring Harbor Laboratory Press).
- Mutoh N.** und Hayashi Y. (1994). Molecular cloning and nucleotide sequencing of *Schizosaccharomyces pombe* homologue of the class II fructose-1,6-bisphosphate aldolase gene. Biochim. Biophys. Acta 1183: 550–552.

8. Literaturverzeichnis

- Nozaki H.**, Matsuzaki M., Takahara M., Misumi O., Kuroiwa H., Hasegawa M., Shin-i T., Kohara Y., Ogasawara N., Kuroiwa T. (2003). The phylogenetic position of red algae revealed by multiple nuclear genes from mitochondria-containing eukaryotes and an alternative hypothesis on the origin of plastids. *J. Mol. Evol.* 56: 485–497.
- Palmer J.D.** (2003). The Symbiotic Birth and Spread of Plastids: How many Times and whodunit? *J. Phycol.* 39: 4–11.
- Patron N.J.**, Rogers M.B., Keeling P.J. (2004). Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot. Cell* 3: 1169–1175.
- Pelzer-Reith B.**, Freund S., Schnarrenberger C., Yatsuki H., Hori K. (1995). The plastid aldolase gene from *Chlamydomonas reinhardtii*: intron/exon organization, evolution, and promoter structure. *Mol. Gen. Genet.* 248: 481–486.
- Petersen J.** (1995). Phylogenetische und molekularbiologische Untersuchungen am Gensystem der plastidären Glycerinaldehyd-3-phosphat-Dehydrogenase (GAPDH) aus Grünalgen und Landpflanzen. Diplomarbeit.
- Petersen J.** (1999). Duplikationen und lateraler Transfer von Genen als Motor der Evolution: Molekulare Studien am GAPDH-Gensystem von Grünalgen, Landpflanzen und Chlorarachniophyta. Dissertation.
- Petersen J.**, Brinkmann H. und Cerff R. (2003). Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. *J. Mol. Evol.* 57: 16–26.
- Petersen J.**, Teich R., Brinkmann H. und Cerff R. (2006a). A “green” phosphoribulokinase in complex algae with red plastids: evidence for a single secondary endosymbiosis leading to haptophytes, cryptophytes, heterokonts and dinoflagellates. *J. Mol. Evol.* 62: 143–157.
- Petersen J.**, Teich R., Becker B., Cerff R. und Brinkmann H. (2006b): The GapA/B gene duplication marks the origin of Streptophyta (charophytes and land plants). *Mol. Biol. Evol.* 23: 1109–1118.
- Pfanzagl B.**, Zenker A., Pittenauer E., Allmaier G., Martinez-Torrecuadrada J., Schmid E. R., De Pedro M. A. und Löffelhardt W. (1996). Primary structure of cyanelle peptidoglycan of *Cyanophora paradoxa*: a prokaryotic cell wall as part of an organelle envelope. *J. Bacteriol.* 178: 332–339.
- Philippe H.**, Lopez P., Brinkmann H., Budin K., Germot A., Laurent J., Moreira D., Müller M. und Le Guyader H. (2000). Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proc. R. Soc. Lond. B* 267: 1213–1221.
- Philippe H.** und Germot A. (2000). Phylogeny of Eukaryotes Based on Ribosomal RNA: Long-Branch Attraction and Models of Sequence Evolution. *Mol. Biol. Evol.* 17(5): 830–834.
- Plaumann M.**, Pelzer-Reith B., Martin W.F. und Schnarrenberger C. (1997). Multiple recruitment of class-I aldolase to chloroplasts and eubacterial origin of eukaryotic class-II aldolases revealed by cDNAs from *Euglena gracilis*. *Curr. Genet.* 31: 430–438.

8. Literaturverzeichnis

- Plaxton W.C.** (1996). The organization and regulation of plant glycolysis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47: 185-214.
- Pohlmeyer K.**, Paap B.K., Soll J. und Wedel N. (1996). CP12: a small nuclear-encoded chloroplast protein provides novel insights into higher-plant GAPDH evolution. *Plant Mol. Biol.* 32: 969–978.
- Qiu Y.L.**, Cho Y., Cox J.C. und Palmer J. D. (1998). The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* 394: 671-674.
- Raines C.A.** (2003). The Calvin cycle revisited. *Photosynth Res* 75: 1-10.
- Robart A.R.** und Zimmerly S. (2005). Group II intron retroelements: function and diversity. *Cytogenetic and Genome Research* 110: 589-597.
- Robbens S.**, Petersen J., Brinkmann H., Rouzé P. und Van de Peer Y.. The ultrasmall green alga *Ostreococcus* unveils a unique regulation of the Calvin cycle. submitted to JME
- Rodríguez-Ezpeleta N.**, Brinkmann H., Burey S.C., Roure B., Burger G., Löffelhardt W., Bohnert H.J., Philippe H. und Lang B.F. (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15:1325–1330.
- Rogers M.** und Keeling P.J. (2004). Lateral transfer and recompartmentalization of Calvin cycle enzymes of plants and algae. *J. Mol. Evol.* 58: 367-375.
- Rombauts S.**, Florquin K., Lescot M., Marchal K, Rouze P. und Van de Peer Y. (2003). Computational Approaches to Identify Promoters and cis-Regulatory Elements in Plant Genomes. *Plant Physiology* 132: 1162–1176.
- Ronimus R.S.** und Morgan H.W. (2003). Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* 1: 199–221.
- Roy S.W.** (2003). Recent evidence for the Exon Theory of Genes. *Genetica* 118: 251-266.
- Roy S.W.** und Gilbert W. (2006). The evolution of spliceosomal introns: patterns, puzzles and progress. *Natur Reviews Genetics* 7: 211-221.
- Sakamoto T.**, Kamiya N., Ueguchi-Tanaka M., Iwahori S. und Matsuoka M. (2001). KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes & Development* 15: 581-590.
- Saldarriaga J.F.**, Taylor F.J.R., Keeling P.J., Cavalier-Smith T. (2001). Dinoflagellate Nuclear SSU rRNA Phylogeny Suggests Multiple Plastid Losses and Replacements. *J. Mol. Evol.* 53: 204-213.
- Sambrook J.**, Fritsch E.F. und Maniatis T. (1989). *Molecular cloning-a laboratory manual* (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).

8. Literaturverzeichnis

- Sánchez Puerta M.V.**, Bachvaroff T.R., Delwiche C.F. (2004). The complete mitochondrial genome sequence of the haptophyte *Emiliana huxleyi* and its relation to heterokonts. *DNA Res.* 11: 1-10.
- Schnarrenberger C.**, Oeser A. und Tolbert N.E. (1973). Two isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Arch. Biochem. Biophys.* 154: 438–448.
- Schnarrenberger C.**, Jacobshagen S., Müller B. und Krüger I. (1990). Evolution of isoenzymes of sugar phosphate metabolism in green algae. *Prog. Clin. Biol. Res.* 344:743-764.
- Schnarrenberger C.**, Pelzer-Reith B., Yatsuki H., Freund S., Jacobshagen S., Hori K. (1994). Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. *Arch. Biochem. Biophys.* 313: 173–178.
- Schnarrenberger C.**, Flechner A. und Martin W. (1995). Enzymatic evidence for a complete oxidative pentose phosphate pathway in chloroplasts and an incomplete pathway in the cytosol of spinach leaves. *Plant Physiol.* 108:609–614.
- Schnepf E.** und Elbraechter M. (1999). Dinophyte plastids and phylogeny: a review. *Grana* 38: 81-97.
- Schwarz-Sommer Z.S.**, Shepherd N., Tacke E., Gierl A., Rohde W., Leclerc L., Mattes M., Berndtgen R., Petersen P.A. und Saedler H. (1987). Influence of transposable elements on the structure and function of the A1 gene of *Zea mays*. *EMBO Journal* 6: 287-294.
- Short J.M.**, Fernandez J.M., Sorge J.A., Huse W.D. (1988) *Nucleic Acids Research* 16: 7583.
- Simon A.**, Glockner G., Felder M., Melkonian M. und Becker B. (2006). EST analysis of the scaly green flagellate *Mesostigma viride* (Streptophyta): Implications for the evolution of green plants (Viridiplantae). *BMC Plant Biology* 6: 2.
- Sparla F.**, Pupillo P. und Trost P. (2002). The C-terminal Extension of Glyceraldehyde-3-phosphate Dehydrogenase Subunit B Acts as an Autoinhibitory Domain Regulated by Thioredoxins and Nicotinamide Adenine Dinucleotide. *The Journal of Biological Chemistry* 277 (47): 44946–44952.
- Sverdlov A.V.**, Rogozin I.B., Babenko V.N. und Koonin E.V. (2004). Reconstruction of ancestral protosplice sites. *Curr. Biol.* 14: 1505–1508.
- Takishita K.**, Ishida K., Maruyama T. (2004). Phylogeny of nuclear-encoded plastid-targeted GAPDH gene supports separate origins for the peridinin- and the fucoxanthin derivative-containing plastids of dinoflagellates. *Protist* 155: 447-458.
- Takishita K.**, Patron N.J., Ishida K.I., Maruyama T., Keeling P.J. (2005). A Transcriptional Fusion of Genes Encoding Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and Enolase in Dinoflagellates. *J. Eukaryot. Microbiol.*, 52(4): 343–348
- Tamoi M.**, Miyazaki T., Fukamizo T., Shigeoka S. (2005). The Calvin cycle in cyanobacteria is regulated by CP12 via the NAD(H)/NADP(H) ratio under light/dark conditions. *Plant J.* 42: 504–513.

8. Literaturverzeichnis

- Tamoi M.**, Nagaoka M., Miyagawa Y., Shigeoka S. (2006). Contribution of fructose-1,6-bisphosphatase and sedoheptulose-1,7-bisphosphatase to the photosynthetic rate and carbon flow in the Calvin cycle in transgenic plants. *Plant Cell Physiol.* 47(3): 380–390.
- Teich R.** (2002). Klonierung und Charakterisierung der Gene für cytosolische Glucose-6-phosphat Isomerase und cytosolisch/plastidäre Isoformen der Glycerinaldehyd-3-phosphat Dehydrogenase aus Haptophyten (*Pavlova lutheri* und *Prymnesium parvum*). Diplomarbeit.
- Teich R.**, Zauner S., Brinkmann H., Petersen J.. Origin and Distribution of Calvin Cycle Fructose and Sedoheptulose Bisphosphatase in Plantae and Complex Algae. A Single Secondary Origin of Complex Red Plastids and Subsequent Propagation via Tertiary Endosymbioses. *Protist in press*.
- Tengs T.**, Dahlberg O.J., Shalchian-Tabrizi K., Klaveness D., Rudi K., Delwiche Ch.F., Jakobsen K.S. (2000). Phylogenetic analysis indicate that the 19'hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. *Mol. Biol. Evol.* 17: 718-729.
- Thein K.** (2003). Molekularbiologische und phylogenetische Charakterisierung Calvin Zyklus spezifischer cDNA-Sequenzen der Rotalge *Chondrus crispus*. Studienarbeit.
- Tolstrup N.**, Rouzé P., Brunak S. (1997). A branch point consensus from *Arabidopsis* found by non-circular analysis allows for better prediction of acceptor sites. *Nucleic Acids Research* 25 (15): 3159–3163
- Turmel M.**, Otis C., De Cambiaire J.C., Pombert J.F., Lemieux C. (2002). The chloroplast genome sequence of *Chlorokybus atmophyticus*: evidence that charophycean green algae from an early diverging lineage adapted to terrestrial life. *J. Phycol.* 38(Suppl. 1): 35–36.
- Unkles S.**, Logsdon J.M., Robison K., Kinghorn J.R. Duncan J.M. (1997). The *tigA* Gene Is a Transcriptional Fusion of Glycolytic Genes Encoding Triose-Phosphate Isomerase and Glyceraldehyde-3-Phosphate Dehydrogenase in Oomycota. *Journal of Bacteriology* 179 (21): 6816-6823.
- van den Hoek Ch.**, Jahns H.M., Mann D.G. (1993). Algen. Thieme Verlag.
- Van de Peer Y.** und De Wachter R. (1997). Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. *J. Mol.Evol.* 45: 619-630.
- Voet D.** und Voet J.G. (1992). Biochemie. VCH Verlagsgesellschaft mbH.
- Wang Y.** und Zhang Z. (2000). Comparative sequence analyses reveal frequent occurrence of short segments containing an abnormally high number of nonrandom base variations in bacterial rRNA genes. *Microbiology* 146: 2845–2854.
- Waller R.F.**, Keeling P.J., van Dooren G.G., McFadden G.I. (2003). Comment on "A green algal apicoplast ancestor". *Science* 301: 49a.
- Watanabe M.M.** und Sasa T. (1991): Major carotenoid composition of an endosymbiont in a green dinoflagellate, *Lepidinium viride*. *J. Phycol.* 27(Sup): 75.

8. Literaturverzeichnis

- Wedel N.**, Soll J. und Paap B.K. (1997). CP12 provides a new mode of light regulation of Calvin cycle activity in higher plants. *Proc. Natl. Acad. Sci. USA* 94: 10479–10484.
- Wedel N.** und Soll J. (1998). Evolutionary conserved light regulation of Calvin cycle activity by NADPH-mediated reversible phosphoribulokinase/CP12/glyceraldehyde-3-phosphate dehydrogenase complex dissociation. *Proc. Natl. Acad. Sci. USA* 95: 9699–9704.
- Weeden N.F.** (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J. Mol. Evol.* 17(3): 133-9.
- Wellman C.H.**, Gray J. (2000). The microfossil record of early land plants. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 355: 717–732.
- Whelan S.** und Goldman N. (2001). A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Mol. Biol. Evol.* 18(5): 691–699.
- Whittaker R.H.** (1959). On the broad classification of organisms. *The Quartely Review of Biology* 34: 210-226.
- Wilson A.P.**, Thelen J.J., Lakritz J., Brown C.R. und Marsh A.E. (2004). The identification of a sequence related to apicomplexan enolase from *Sarcocystis neurona*. *Parasitol. Res.* 94: 354–360.
- Woese C.R.** (1987). Bacterial Evolution. *Microbiological Reviews* 51 (2): 221-271.
- Woese C.R.**, Kandler O., Wheelis M.L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* 87: 4576-4579.
- Wood W.B.** (1966). *Journal of Molecular Biology* 16: 118.
- Yano H.**, Wong J.H., Lee Y.M., Cho M.J., Buchanan B.B. (2001). A strategy for the identification of proteins targeted by thioredoxin. *Proc. Natl. Acad. Sci. USA* 98: 4794-4799.
- Yoon H.S.**, Hackett J.D., Van Dolah F.M., Nosenko T., Lidie K.L., Bhattacharya D. (2005). Tertiary Endosymbiosis Driven genome Evolution in Dinoflagellate Algae. *Mol. Biol. Evol.* 22(5):1299–1308.
- Zhang Z.**, Green B.R., Cavalier-Smith T. (1999). Single gene circles in dinoflagellate chloroplast genomes. *Nature* 400: 155.
- Zuckerkindl E.**, Jones R.T. und Pauling L. (1960). A Comparison of Animal Hemoglobins by Tryptic Peptide Pattern Analysis. *Proc. Nat. Acad. Sci. USA* 46 (10): 1349-60.

9. Anhang

9.1 Accession Numbers

9.1.1 Accession Numbers Teil A (cDNA Klone)

<i>Mesostigma viride</i>	GapA	DQ270265
<i>Marchantia polymorpha</i>	GapB	AJ246027

9.1.2 Accession Numbers Teil A (PCR mit genomischer DNA)

<i>Mesostigma viride</i>	GapA	DQ873397
	GapB	DQ873398
	GapC	DQ873399
<i>Chlorokybus atmophyticus</i>	GapA	DQ873408
<i>Chara vulgaris</i>	GapA	DQ873407
<i>Coleochaete scutata</i>	GapC	DQ873409
<i>Marchantia polymorpha</i>	GapA	DQ873393
	GapB	DQ873394
	GapC	DQ873395
	GapCp	DQ873396

9.1.3 Accession-Numbers Teil B

Phosphoribulokinase (PRK)

<i>Prymnesium parvum</i> (genomischer Klon)	AY772252
<i>Pavlova lutheri</i> (cDNA Klon)	AY772245
<i>Guillardia theta</i> (cDNA Klon)	AY772246
<i>Hanusia phi</i> (cDNA Klon)	AY772251
<i>Bigelowiella natans</i> (genomische PCR)	AY772253

9. Anhang

Fructose-1,6-bisphosphatase (FBP)

<i>Lingulodinium polyedrum</i> (cDNA Klon)	DQ508159
--	----------

Sedoheptulose-1,7-bisphosphatase (SBP)

<i>Prymnesium parvum</i> (genomischer Klon)	DQ508152
---	----------

<i>Lingulodinium polyedrum</i> (cDNA Klon)	DQ508153
--	----------

<i>Mesostigma viride</i> (genomische PCR)	DQ873402
---	----------

<i>Guillardia theta</i> (genomische PCR)	DQ508165
--	----------

<i>Chondrus crispus</i> (genomische PCR)	DQ873403
--	----------

Die Sequenzen sind in der „DDBJ/EMBL/GenBank International Nucleotide Sequence Database“ unter den angegebenen Accession Nummern hinterlegt. Die Hinterlegung erfolgte vertraulich, die Sperrfrist läuft am 31.12.2009 aus.

9. Anhang

9.2 Sequenzen

9.2.1 *Pavlova lutheri* Triosephosphat Isomerase (TPI; cDNA Klon)

													M	G	G	G	R	K	M
AA	nt	gcc	ccg	ctt	cct	gcc	cga	acg	cat	cgc	tcg	acc	ATG	GGT	GGC	GGA	CGC	AAG	ATG
													36			45			54
10																			
I	V	G	G	N	W	K	C	N	L	D	Q	A	K	A	A	A	L	V	E
ATC	GTT	GGC	GGC	AAC	TGG	AAG	TGC	AAC	CTG	GAC	CAG	GCG	AAG	GCG	GCG	GCC	CTC	GTC	GAG
		63			72			81			90			99			108		
40																			
Q	L	N	G	M	E	V	G	N	C	E	V	V	V	A	P	V	F	L	H
CAG	CTC	AAC	GGC	ATG	GAG	GTT	GGC	AAC	TGC	GAG	GTC	GTC	GTC	GCG	CCG	GTC	TTT	CTG	CAC
		123			132			141			150			159			168		
70																			
I	P	M	V	M	N	T	L	K	K	D	K	A	L	V	S	A	Q	N	C
ATC	CCG	ATG	GTG	ATG	AAC	ACG	CTC	AAG	AAG	GAC	AAG	GCG	CTC	GTC	TCG	GCG	CAG	AAC	TGC
		183			192			201			210			219			228		
90																			
S	F	T	G	M	G	A	Y	T	G	D	I	S	A	E	H	L	V	D	M
TCG	TTC	ACG	GGC	ATG	GGT	GCG	TAC	ACG	GGC	GAC	ATC	TCG	GCC	GAG	CAC	CTC	GTC	GAC	ATG
		243			252			261			270			279			288		
100																			
G	I	Q	W	V	I	L	G	H	S	E	R	R	Q	Y	F	K	E	D	D
GGC	ATC	CAG	TGG	GTC	ATC	CTT	GGC	CAC	TCC	GAG	CGC	CGC	CAG	TAC	TTC	AAG	GAG	GAC	GAT
		303			312			321			330			339			348		
110																			
E	M	L	A	K	K	L	E	Y	A	L	S	K	G	L	K	V	I	Y	C
GAG	ATG	CTC	GCC	AAG	AAG	CTC	GAG	TAC	GCG	CTC	TCG	AAG	GGG	CTC	AAG	GTG	ATC	TAC	TGC
		363			372			381			390			399			408		
130																			
I	G	E	Q	K	E	Q	R	E	N	G	T	T	M	D	V	C	K	K	Q
ATC	GGC	GAG	CAG	AAG	GAG	CAG	CGC	GAG	AAC	GGC	ACG	ACG	ATG	GAC	GTC	TGC	AAG	AAG	CAG
		423			432			441			450			459			468		
140																			
160																			
L	E	G	V	A	K	M	L	N	P	E	T	V	V	I	A	Y	E	P	V
CTC	GAG	GGC	GTC	GCC	AAG	ATG	CTC	AAC	CCC	GAG	ACG	GTT	GTG	ATC	GCG	TAC	GAG	CCC	GTT
		483			492			501			510			519			528		
170																			
180																			
W	A	I	G	T	G	L	V	A	T	P	E	Q	A	Q	D	T	Q	K	S
TGG	GCC	ATC	GGC	ACC	GGC	CTC	GTT	GCC	ACG	CCG	GAG	CAG	GCG	CAG	GAC	ACG	CAG	AAG	TCG
		543			552			561			570			579			588		
190																			
I	R	E	Y	I	A	S	A	A	S	P	D	V	A	A	K	I	R	I	Q
ATC	CGC	GAG	TAC	ATC	GCG	TCG	GCG	GCA	TCG	CCC	GAC	GTG	GCC	GCC	AAG	ATC	CGC	ATC	CAG
		603			612			621			630			639			648		

9. Anhang

210

Y	G	G	S	V	N	G	K	N	C	A	E	L	A	A	L	E	D	I	D
TAC	GGC	GGC	TCG	GTC	AAC	GGC	AAG	AAC	TGC	GCG	GAG	CTC	GCG	GCG	CTC	GAG	GAC	ATT	GAC
		663			672			681			690			699			708		

230

240

G	F	L	V	G	G	A	S	L	K	P	E	F	T	D	I	V	K	A	C
GGC	TTC	CTC	GTC	GGC	GGC	GCG	TCG	CTC	AAG	CCC	GAG	TTC	ACC	GAC	ATT	GTC	AAG	GCT	TGC
		723			732			741			750			759			768		

S	A	.																	
TCG	GCG	TAA	gtg	gga	cgc	ggc	tca	gcg	gcg	ggc	gcc	gcg	tgc	act	ggc	ggc	acg	gag	atg
		783			792			801			810			819			828		

cgg	gcg	gcc	gcc	agc	tgt	aag	ggg	tgc	ggc	gtg	ccg	ccg	ccg	ggc	gcg	cgg	cca	cgc	gct
		843			852			861			870			879			888		

agg	ccc	atc	gcc	gtg	cac	agc	ctt	tgc	acg	tcg	tgc	cat	ctc	tca	act	caa	ctg	gtt	tcg
		903			912			921			930			939			948		

aaa	tgc	tgt	gga	cag	ggc	gcg	ccg	gaa	aga	gcg	can	aga	ggt	cgg	acg	ttc	acg	gtc	tac
		963			972			981			990			999			1008		

gtc	ccc	agt	cgg	tcc	caa	ttt	cct	ncg	agg	cgg	gaa	aaa	gct	gtt	cgg	tca	aat	cga	ccc
		1023			1032			1041			1050			1059			1068		

cgc	cgt	cct	ccc	ttt	tcc	ttc	ggg	cat	gca	cga	nct	tct	tg
		1083			1092			1101			1110		

9. Anhang

9.2.2 *Pavlova lutheri* Enolase (cDNA Klon)

AA								M	A	T	I	A	K	R	T	R	N	V
nt	tcg	cga	acc	ccg	gct	ccg	ATG	GCC	ACC	ATC	GCC	AAG	CGC	ACG	CGG	CGC	AAC	GTG
			9			18			27			36			45			54
																		20
P	A	K	I	T	K	V	H	A	R	E	I	Y	D	S	R	G	N	P
CCC	GCG	AAG	ATC	ACC	AAG	GTG	CAC	GCG	CGC	GAG	ATC	TAC	GAC	AGC	CGT	GGC	AAC	CCA
		63			72			81			90			99			108	
																		40
L	E	V	E	L	T	T	P	D	G	T	F	R	A	D	V	P	S	G
CTC	GAG	GTG	GAG	CTC	ACG	ACG	CCG	GAC	GGC	ACG	TTC	CGC	GCG	GAC	GTG	CCG	TCC	GGC
		123			132			141			150			159			168	
																		60
S	T	G	D	K	E	A	H	E	L	R	D	G	G	A	R	L	S	G
TCG	ACG	GGC	GAC	AAG	GAG	GCG	CAC	GAG	CTG	CGC	GAC	GGC	GGC	GCG	CGC	CTC	TCG	GGC
		183			192			201			210			219			228	
								70										80
G	V	A	K	A	V	T	N	V	I	D	V	I	G	P	A	I	L	G
GGC	GTG	GCG	AAG	GCT	GTG	ACG	AAC	GTG	ATC	GAC	GTG	ATC	GGC	CCC	GCC	ATC	CTG	GGC
		243			252			261			270			279			288	
								90										100
D	A	I	D	I	A	S	L	D	A	Q	L	C	A	L	D	G	T	D
GAC	GCG	ATC	GAC	ATC	GCC	TCG	CTC	GAC	GCG	CAG	CTC	TGC	GCG	CTC	GAC	GGC	ACA	ACG
		303			312			321			330			339			348	
		110											120					
K	S	K	L	G	A	N	A	I	L	G	V	S	M	A	A	C	R	A
AAG	TCC	AAG	CTC	GGC	GCC	AAC	GCC	ATC	CTG	GGC	GTC	TCG	ATG	GCG	GCG	TGC	CGC	GCC
		363			372			381			390			399			408	
												140						
A	A	A	A	G	E	P	L	Y	V	F	L	N	R	L	A	G	S	P
GCT	GCG	GCG	GCG	GGT	GAG	CCG	CTG	TAT	GTC	TTT	CTC	AAC	CGG	CTC	GCC	GGC	TCG	CCC
		423			432			441			450			459			468	
		150										160						
M	V	M	P	M	P	C	F	N	L	V	N	G	G	V	H	A	G	N
ATG	GTC	ATG	CCG	ATG	CCG	TGC	TTC	AAC	CTT	GTG	AAC	GGT	GGC	GTG	CAC	GCG	GGC	AAC
		483			492			501			510			519			528	
		170										180						
L	P	F	Q	E	F	F	A											

9. Anhang

																				240					
A	L	K	L	I	V	S	A	V	E	R	A	G	L	T	G	K	V	T	L						
GCG	CTC	AAG	CTG	ATC	GTC	AGC	GCG	GTC	GAA	CGG	GCC	GGC	CTC	ACT	GGC	AAG	GTC	ACG	CTC						
		723						732				741				750				759				768	
250										260															
G	C	D	P	A	A	S	E	M	Y	D	K	D	R	G	V	Y	N	L	D						
GGC	TGC	GAC	CCA	GCG	GCG	TCC	GAG	ATG	TAT	GAC	AAG	GAT	CGC	GGT	GTA	TAC	AAC	CTC	GAC						
		783						792				801				810				819				828	
F	K	K	P	A	A	E	Q	Q	P	E	N	I	M	K	G	D	D	L	V						
TTC	AAG	AAG	CCG	GCG	GCC	GAG	CAG	CAG	CCC	GAA	AAC	ATC	ATG	AAG	GGA	GAC	GAC	CTC	GTC						
		843						852				861				870				879				888	

9. Anhang

ccg cag gcg ctg tgc act tgt ggc gct gcc acg gct gcc cga cgt cag aac acc aag tcg
1503 1512 1521 1530 1539 1548

													Aminosäurenummer
													Aminosäuresequenz
tct	gca	tcc	aaa	aaa	aaa	aaa	aaa	aaa	aaa	aaa	aaa	a	Nukleotidsequenz
			1563			1572			1581			1590	Nukleotidnummer

9. Anhang

9.2.3 *Pavlova lutheri* Fructose-1,6-bisphosphat Aldolase (FBA Klasse I; cDNA Klon)

AA	gag		acc	tca	gcg	cac	caa	gca	acA	M	S	A	A	A	A	T	R	P	E
nt										TGT	CCG	CCG	CCG	CCG	CCG	CCA	CGC	GTC	CCG
				9			18			27			36			45			54
	10																		
F	C	K	F	A	A	E	L	K	A	T	A	Q	A	I	A	A	P	G	K
AGT	TCT	GCA	AGT	TCG	CCG	CGG	AGC	TGA	AGG	CGA	CCG	CCC	AGG	CGA	TCG	CCG	CGC	CCG	GCA
		63			72			81			90			99			108		
	30																		
G	I	L	A	A	D	E	S	G	G	T	I	G	K	R	F	D	S	I	K
AGG	GCA	TCC	TCG	CCG	CCG	ACG	AGT	CGG	GCG	GCA	CGA	TCG	GCA	AGC	GCT	TCG	ACT	CGA	TCA
		123			132			141			150			159			168		
	50																		
V	E	N	T	Q	E	N	R	R	G	Y	R	E	L	L	F	T	T	E	G
AGG	TCG	AGA	ACA	CGC	AGG	AGA	ACC	GGC	GCG	GGT	ATC	GCG	AGC	TGC	TTT	TCA	CCA	CCG	AGG
		183			192			201			210			219			228		
	70																		
L	G	E	Y	I	S	G	C	I	L	F	E	E	T	L	F	D	V	A	A
GGC	TCG	GCG	AGT	ACA	TCT	CCG	GGT	GCA	TCC	TGT	TCG	AGG	AGA	CGC	TCT	TCG	ACG	TTG	CGG
		243			252			261			270			279			288		
	90																		
D	G	T	T	K	L	V	D	L	L	H	K	Q	G	I	I	P	G	I	K
CCG	ACG	GCA	CGA	CCA	AGC	TCG	TCG	ACC	TGC	TCC	ACA	AGC	AGG	GCA	TCA	TCC	CCG	GTA	TCA
		303			312			321			330			339			348		
	110																		
V	D	K	G	V	Q	P	L	P	G	T	S	G	E	T	V	T	Q	G	M
AGG	TCG	ACA	AGG	GCG	TGC	AGC	CGC	TGC	CTG	GCA	CGT	CGG	GCG	AGA	CGG	TCA	CGC	AAG	GCA
		363			372			381			390			399			408		
	130																		
D	D	L	G	K	R	C	A	K	Y	Y	E	A	G	A	R	F	A	K	W
TGG	ACG	ACC	TGG	GCA	AGC	GCT	GCG	CCA	AGT	ACT	ACG	AGG	CGG	GCG	CGC	GCT	TCG	CCA	AGT
		423			432			441			450			459			468		
	140																		
R	A	V	L	N	I	N	D	S	T	G	A	T	P	S	D	L	A	I	K
GGC	GCG	CGG	TCC	TCA	ACA	TCA	ACG	ACT	CGA	CGG	GCG	CGA	CCC	CGT	CCG	ACC	TCG	CGA	TTA
		483			492			501			510			519			528		
	170																		
Q	N	A	E	T	L	A	R	Y	A	V	I	C	Q	M	N	G	L	V	P
AGC	AGA	ACG	CCG	AGA	CGC	TCG	CGC	GCT	ACG	CGG	TCA	TCT	GCC	AGA	TGA	ACG	GCC	TCG	TGC
		543			552			561			570			579			588		
	190																		
I	V	E	P	E	V	L	M	D	G	A	H	D	I	D	M	A	Q	R	V
CGA	TCG	TCG	AGC	CCG	AGG	TGC	TGA	TGG	ACG	GCG	CGC	ACG	ACA	TTG	ACA	TGG	CGC	AGC	GCG
		603			612			621			630			639			648		
	210																		
T	E	K	V	I	A	A	C	Y	K	E	L	S	D						

9. Anhang

230

G	T	L	L	K	P	N	M	V	R	S	G	E	A	A	P	A	D	K	K
AGG	GAA	CTC	TGC	TCA	AGC	CCA	ACA	TGG	TCC	GCA	GCG	GCG	AGG	CGG	CGC	CCG	CCG	ACA	AGA
	723				732			741			750			759			768		

250

A	T	A	A	Q	I	A	A	A	T	V	T	C	L	Q	R	T	V	P	P
AGG	CGA	CGG	CGG	CGC	AGA	TCG	CCG	CGG	CGA	CGG	TCA	CGT	GCC	TGC	AGC	GCA	CGG	TGC	CGC
	783				792			801			810			819			828		

260 **270**

A	V	P	G	I	T	F	L	S	G	G	M	S	E	E	E	A	T	V	V
CCG	CCG	TGC	CGG	GCA	TCA	CCT	TCC	TCT	CCG	GCG	GCA	TGT	CGG	AGG	AGG	AGG	CGA	CCG	TCG
	843				852			861			870			879			888		

280 **290**

L	N	E	M	N	K	L	A	T	V	K	P	W	A	L	T	F	S	Y	G
TGC	TCA	ACG	AGA	TGA	ACA	AGC	TCG	CCA	CCG	TCA	AGC	CGT	GGG	CGC	TCA	CCT	TCT	CGT	ACG
	903				912			921			930			939			948		

300 **310**

R	A	L	Q	Q	S	T	L	K	V	W	M	G	K	K	E	N	E	G	A
GCC	GCG	CGC	TGC	AGC	AGT	CGA	CAC	TCA	AGG	TGT	GGA	TGG	GCA	AGA	AGG	AGA	ACG	AAG	GCG
	963				972			981			990			999			1008		

330

A	K	K	E	L	I	V	R	A	K	A	N	S	L	A	V	Q	G	K	Y
CGG	CGA	AGA	AGG	AGC	TGA	TCG	TGC	GCG	CCA	AGG	CCA	ACT	CGC	TCG	CCG	TCC	AGG	GCA	AGT
	1023				1032			1041			1050			1059			1068		

N	G	E	A	A	G	G	A	A	A	A	T	S	T	F	V	P	S	Y	S
ACA	ACG	GCG	AGG	CGG	CCG	GCG	GCG	CGG	CCG	CGG	CGA	CGT	CGA	CGT	TCG	TGC	CGT	CGT	ACT
	1083				1092			1101			1110			1119			1128		

Y	.																		
CGT	ACT	AGg	cgc	gcg	gcg	gcc	gcc	act											
	1143				1152			1161											

Aminosäuresequenz
Nukleotidsequenz
Nukleotidnummer

9. Anhang

9.2.4 Phosphoribulokinase (PRK) *Hanusia phi* und *Guillardia theta*

P356						M	G	T	P	L	A	S	R	S	S	F	A																		
CTC	TCG	TCT	GCT	GCT	CAC	TTC	GGA	ACC	CCC	CTC	GCT	TCC	CGC	AGC	ACC	TTC	GCC																		
9						18						27						36						45						54					
S	P	V	A	L	R	V	T	N	D	K	E	M	D	V	A	I	Q	Q	I																
S	P	V	S	L	R	V	T	N	D	K	E	M	D	L	A	I	Q	Q	I																
AGC	CCC	GTC	TCC	CTC	CGT	GTC	ACC	AAC	GAC	AAG	GAG	ATG	GAT	CTT	GCT	ATC	CAG	CAG	ATC																
63						72						81						90						99						108					
S	M	S	L	A	S	G	Q	K	P																										
S	M	S	L	A	S	G	Q	K	P	>Intron 51-0<																									
AGC	ATG	TCC	CTT	GCC	AGT	GGA	CAG	AAG	CCT	gta	aga	tgc	gca	ggg	gcc	tcg	agc	atg	ccc																
123						132						141						150						159						168					
						V	V	I	G	L	A	A	D	S	G	C	G	K	S																
						V	V	I	G	L	A	A	D	S	G	C	G	K	S																
ttg	ctg	acg	gct	ctc	tag	GTT	GTC	ATC	GGC	CTC	GCT	GCT	GAC	TCT	GGC	TGC	GGA	AAG	TCG																
183						192						201						210						219						228					
T	F	M	R	R	V	T	A	C	F	G	G	A	C	K	L	N	P	I	G																
T	F	M	R	R	V	T	A	C	F	G	G	A	S	K	L	N	P	I	G																
ACC	TTC	ATG	CGC	CGT	GTG	ACC	GCT	TGC	TTC	GGC	GGA	GCC	AGC	AAG	CTC	AAC	CCC	ATT	GGC																
243						252						261						270						279						288					
R	E	T	N	T	L	I	S	D	M	T	T	V	I	C	L	D	D	Y	H																
R	E	T	N	T	L	I	S	D	M	T	T	V	I	C	L	D	D	Y	H																
CGT	GAG	ACC	AAC	ACC	CTC	ATC	TCC	GAC	ATG	ACC	ACC	GTT	ATC	TGC	CTC	GAT	GAC	TAC	CAT																
303						312						321						330						339						348					
L	N	D	R	Q	G	R	K	K	T	G																									
L	N	D	R	Q	G	R	K	K	T	G	>Intron 116-1<																								
CTT	AAC	GAC	AGA	CAG	GGC	CGC	AAG	AAG	ACC	Ggt	ggg	tgg	aga	aat	gtg	ttg	gtt	tcc	aca																
363						372						381						390						399						408					
						L	T	A	L	D	P	R	E	N	N	F	D	L																	
						L	T	A	L	D	P	R	E	N	N	F	D	L																	
gga	tct	gac	tcc	ctg	gac	agG	CCT	CAC	TGC	CCT	GGA	CCC	GCG	TGA	GAA	CAA	TTT	CGA	CCT																
423						432						441						450						459						468					
M	Y	E	Q	V	K	A	L	K	E	G	K	K	V	M	K	P	I	Y	N																
M	Y	E	Q	V	K	A	L	K	E	G	K	K	L	M	K	P	I	Y	N																
CAT	GTA	CGA	GCA	GGT	TAA	GGC	CTT	GAA	GGA	GGG	CAA	GAA	GAT	CAT	GAA	GCC	CAT	CTA	CAA																
483						492						501						510						519						528					
H	V	N	G	T	L	D	E	A	E	E	I	T	P	T	P	I	I	I	F																
H	V	N	G	T	L	D	E	E	E	E	I	T	P	T	P	I	I	I	F																
CCA	CGT	CAA	CGG	AAC	TCT	CGA	TGA	GCC	CGA	GGA	GAT	CAC	CCC	CAC	CCC	CAT	TAT	CAT	CTT																
543						552						561						570						579						588					
E																	G	L	H	P															
E	>Intron 171-0<																G	L	H	P															
CGA	Ggt	acg	tgc	ctt	gcc	gtt	gcc	cag	cag	gcc	ccc	tga	cga	ctt	tcc	agG	GTC	TCC	ACC																
603						612						621						630						639						648					

9. Anhang

F	Y	D	K	R	V	E	E	L	M	D	F	K	I	Y	V	D	I	T	P
F	Y	D	K	R	V	E	E	L	M	D	F	K	I	Y	V	D	I	T	P
CCT	TCT	ACG	ACA	AGC	GTG	TCG	AGG	AGC	TCA	TGG	ACT	TCA	AGA	TCT	ACG	TTG	ACA	TCA	CCC
		663			672			681			690			699			708		
E	V	K	F	N	W	K	V	Q	R	D	H	E	E	R	G	H	S	I	E
E	V	K	F	N	W	K	V	Q	R	D	H	E	E	R	G	H	S	I	E
CCG	AGG	TCA	AGT	TCA	ACT	GGA	AGG	TTC	AGC	GCG	ACC	ACG	AGG	AGC	GCG	GCC	ACT	CCA	TCG
		723			732			741			750			759			768		
S	I	K	Q	Q	I	E	A	R	K	P	D	F	D	A	Y	I	D	P	Q
S	I	K	Q	Q	I	E	A	R	K	P	D	F	D	A	Y	I	D	P	Q
AGT	CCA	TCA	AGC	AGC	AGA	TCG	AGG	CCC	GCA	AGC	CTG	ACT	TCG	ATG	CTT	ACA	TTG	ATC	CCC
		783			792			801			810			819			828		
K	N	K	A	D	C	V	I	Q	V	L	P	T	N	L	V	A	N	D	K
K	N	K	A	D	C	V	I	Q	V	L	P	T	G	L	A	A	N	D	K
AGA	AGA	ACA	AGG	CCG	ACT	GCG	TCA	TCC	AGG	TTC	TCC	CCA	CCC	AGC	TTG	CCG	CCA	ACG	ACA
		843			852			861			870			879			888		
T	H	L																	
T	H	L																	
AGA	CCC	ACC	TC																
		903																	
P337																			

Hanusia phi (cDNA-Klon)

Guillardia theta (genomischer PCR Klon)

nt-Sequenz *G.theta* (genomischer PCR Klon)

nt-Nummer

Verglichen ist der Bereich zwischen den Primern P356 und P337. Unterschiede in der genomischen Sequenz von *Guillardia theta* zu der cDNA Sequenz von *Hanusia phi* sind durch **rote** (Nukleotid Ebene) oder **rot unterlegte** Buchstaben (Aminosäure Ebene) hervorgehoben.

9. Anhang

9.2.5 *Prymnesium parvum* Fructose-1,6-bisphosphatase (FBP; RT-PCR Klon)

AA nt	P090												130					
	V	V	F	D	P	L	D	G	S	S	N	I	D	A	G	I	S	T
	GTC	GTG	TTC	GAC	CCC	CTC	GAT	GGC	AGC	TCG	AAC	ATC	GAC	GCC	GGC	ATC	TCC	ACC
			9			18			27			36			45			54
140																		
G	S	I	F	G	I	Y	D	P	N	D	E	C	S	L	E	D	M	D
GGC	TCC	ATC	TTC	GGC	ATC	TAC	GAC	CCC	AAC	GAC	GAG	TGC	AGC	CTG	GAG	GAC	ATG	GAC
		63			72			81			90			99			108	
150																		
P	E	Q	M	M	K	N	C	V	L	N	V	C	Q	P	G	T	S	L
CCT	GAG	CAG	ATG	ATG	AAG	AAC	TGC	GTG	CTG	AAC	GTG	TGC	CAA	CCA	GGC	ACC	TCG	CTG
		123			132			141			150			159			168	
170																		
G	A	G	Y	C	L	Y	S	S	S	T	V	L	V	L	T	L	G	N
GGC	GCG	GGC	TAC	TGC	CTG	TAC	TCC	TCG	TCC	ACG	GTG	CTG	GTG	CTG	ACG	CTC	GGC	AAC
		183			192			201			210			219			228	
190																		
V	N	G	F	T	Y	D	S	N	V	G	E	F	I	L	S	H	P	N
GTC	AAC	GGC	TTC	ACC	TAC	GAC	TCC	AAC	GTC	GGC	GAG	TTC	ATC	CTG	AGC	CAC	CCC	AAC
		243			252			261			270			279			288	
210																		
Q	I	P	D	K	G	A	I	Y	S	F	N	E	G	N	Y	D	G	W
CAG	ATC	CCC	GAC	AAG	GGC	GCG	ATC	TAC	TCC	TTC	AAC	GAG	GGC	AAC	TAC	GAT	GGC	TGG
		303			312			321			330			339			348	
230																		
E	G	L	K	K	Y	M	D	S	L	K	T	K	D	N	I	N	G	K
GAG	GGC	CTG	AAG	AAG	TAC	ATG	GAC	AGC	CTG	AAG	ACG	AAG	GAC	AAC	ATC	AAC	GGG	AAG
		363			372			381			390			399			408	
250																		
Y	S	S	R	Y	I	G	S	L	V	G	D	F	H	R	T	M	L	Y
TAC	TCA	TCC	CGC	TAC	ATC	GGC	AGC	CTG	GTG	GGC	GAC	TTC	CAC	CGC	ACG	ATG	CTG	TAC
		423			432			441			450			459			468	
270																		
G	I	Y	G	Y	P	G	D	V	K	N	K	N	G	K	L	R	L	L
GGC	ATC	TAC	GGC	TAC	CCG	GGC	GAC	GTG	AAG	AAC	AAG	AAC	GGG	AAG	CTG	CGG	CTG	CTG
		483			492			501			510			519			528	
P091																		
E	C	A	P	M	S	M	I	A	E	Q	A	G	G	Q	A			
GAG	TGC	GCG	CCG	ATG	TCC	ATG	ATC	GCG	GAA	CAA	GCG	GGC	GGG	CAG	GC			
		543			552			561			570			579				

9. Anhang

9.3 Abbildungen

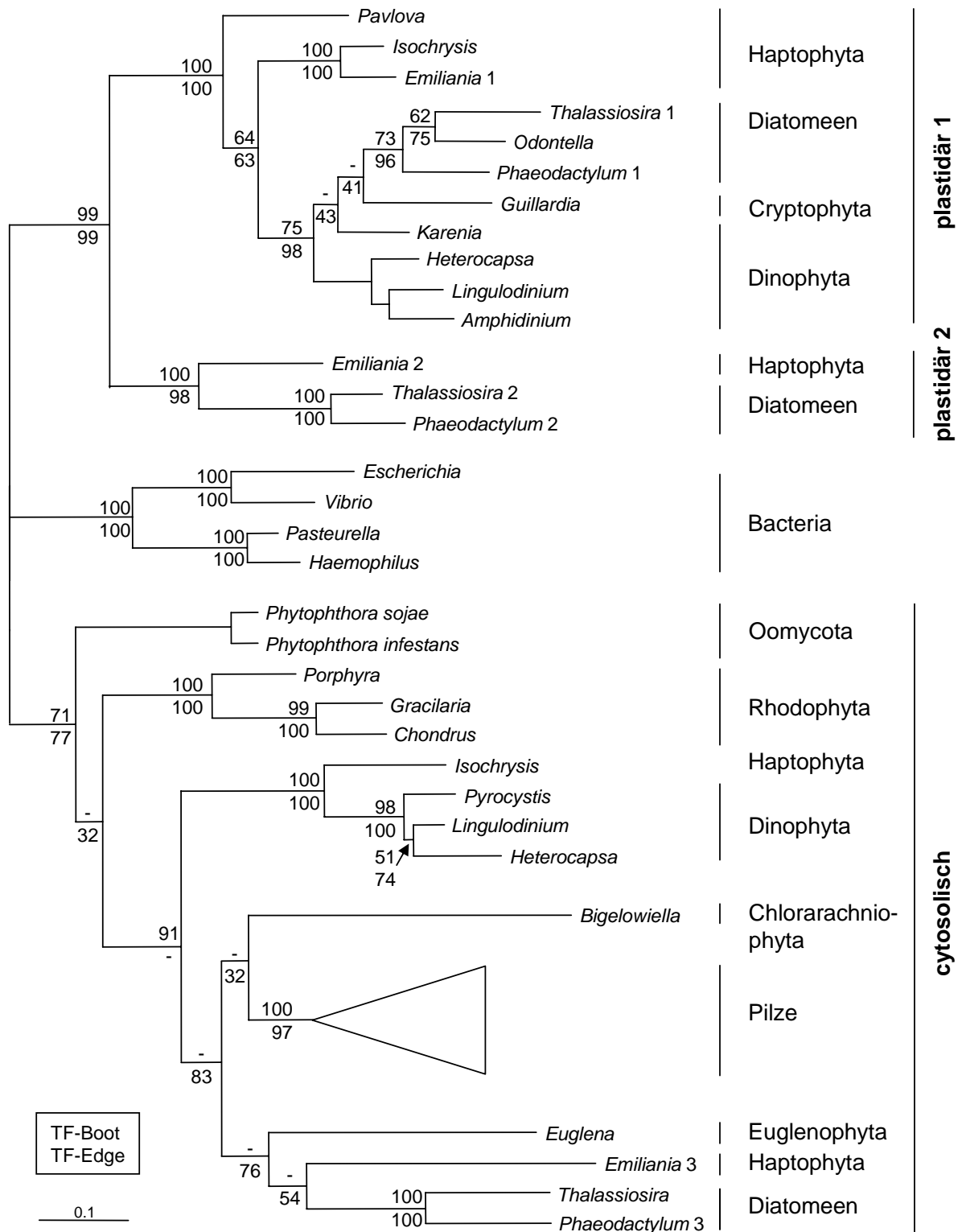


Abbildung 22: Phylogenetische Analyse der FBA Klasse II (TypA)

Maximum Likelihood Analyse (Treefinder) mit WAG-Matrix basierend auf 43 FBA Klasse II Sequenzen. Die statistische Unterstützung der Äste wurde durch je 100 Bootstrap- (oben) und Edgesupport-Analysen (unten) ermittelt.

9. Anhang

	GapA										
	3-0	7-1	7-2	13-2	15-0	19-2	24-0	40-0	41-0	70-0	
GapA											Arabidopsis 1
											Arabidopsis 2
											Pisum
											Medicago
											Oryza
											Zea
											Marchantia
											Chara
											Chlorokybus
											Mesostigma
GapB											Arabidopsis
											Pisum
											Medicago
											Oryza
											Zea
											Marchantia
											Chlorokybus
											Mesostigma
											Chlamydomonas
											Volvox
GapC											Arabidopsis 1
											Arabidopsis 2
											Pisum
											Zea 4
											Marchantia
											Coleochaete
											Mesostigma
											Chlamydomonas
											Volvox
											Arabidopsis 1
GapCp											Arabidopsis 2
											Capsicum
											Pinus
											Marchantia
											Homo
											Gallus
											Xenopus
											Danio
											Takifugu
											Drosophila
GapC											Anopheles
											Caenorhabditis 1&4
											Caenorhabditis 2&3
											Strongylocentrotus
											Branchiostoma
											Oikopleura
											Ciona
											Ustilago
											Podospira
											Curvularia
GapC											Aspergillus
											-Mensch
											-Vogel
											-Frosch
											-Fisch
											-Insekt
											-Insekt
											-Rundwurm
											-Rundwurm
											-Seeigel
Pilze											-Lanzettfischchen
											-Manteltier
											-Seescheide
											-Basidiomycota
											-Ascomycota
											-Ascomycota
											-Ascomycota
											-Ascomycota
											-Ascomycota
											-Ascomycota

Abbildung 23: GAPDH Intronmatrix

S. pombe: kein Intron; *S. cerevisiae*: kein Intron; *Candida albicans*: kein Intron; *Colletotrichum gleosporioides*: 41-0; *Cryphonectria parasitica*: 41-0, 269-1; *Cochliobolus heterostrophus*: 3-0, 19-2, 41-0, 269-1; *Schizophyllum commune*: **9-1, 32-2, 97-0, 226-2**; *Phanerochaete chrysosporium*: **1-0, 7-2, 9-1, 32-2, 41-0, 227-0**; *Agaricus bisporus*: **1-0, 7-2, 9-1, 32-2, 41-0, 97-0, 104-2, 144-0, 226-2**; *Cyanophora paradoxa* GapA: **112-1, 136-0, 170-1, 191-2, 232-1** (persönliche Mitteilung R. Cerff); *Cyanophora paradoxa* GapC: **9-1, 30-0, 44-0, 60-0, 82-0, 124-1, 170-1, 277-0, 310-2** (persönliche Mitteilung R. Cerff); *Galdieria sulphuraria* GapA: 2 Introns im TP; *Cyanidioschyzon merolae* GapA: kein Intron; *Cyanidioschyzon merolae* GapC: kein Intron; *Chondrus crispus* GapA: 1 Intron im TP; *Euglena garcilis* GapA: 1 Intron im TP, 213-0, 4 nicht-spleißosomale Introns; *Euglena gracilis* GapC: 44-0, 3 nicht-spleißosomale Introns. Dargestellt sind nur Introns ab Position +1; Introns, die in der Abbildung nicht auftauchen, sind fett dargestellt. Ein hellgrauer Rahmen markiert Intronpositionen, die in GapC Genen von Grünalgen, Landpflanzen, Tieren und Pilzen identisch sind; dunkelgraue Rahmen markieren identische Positionen in GapA und GapC Genen.

9. Anhang

9.4 Abkürzungen

°C	Grad Celsius
% (v/v)	Volumenprozent
% (w/v)	Gewichtprozent
6-PGDH	6-Phosphoglycerat Dehydrogenase
AA, AS	Aminosäure
Abb.	Abbildung
Acc.-Nr.	Accession Number
ATP	Adenosintriphosphat
ATP-PFK	ATP-abhängige Phosphofructokinase
bp, kb	Basenpaare, Kilobasenpaare
c	Konzentration (in mg/ml)
cDNA	copy Desoxyribonukleinsäure
dNTP	Desoxyribonukleosidtriphosphat
DTT	Dithiotreitol
EDTA	Ethylendiamid-N, N, N', N'-tetraacetat
ENO	Enolase (2-Phospho-D-Glycerat Hydrolase)
EST	expressed sequence tag
EtOH	Ethanol
FBAI	Fructose-1,6-bisphosphat Aldolase Klasse I
FBP	Fructose-1,6-bisphosphatase
g	Gramm, Erdbeschleunigung
G6PDH	Glucose-6-phosphat Dehydrogenase
GapAB	plastidäre GAPDH (Calvin-Zyklus)
GAPDH	Glycerinaldehyd-3-Phosphat Dehydrogenase
GapC	cytosolische GAPDH (Glykolyse)
GapCp	plastidäre GapC
h	Stunde(n)
IPTG	Isopropyl-β-D-thiogalactopyranosid
l	Liter
μ, m	mikro, milli
min	Minute(n)
mRNA	messenger Ribonukleinsäure
n	nano
NAD(P)H	Nicotinamid-Adenin-Dinukleotid
nt	Nukleotid
OD	optische Dicht bei der Wellenlänge
p	pico
PCI	Phenol/Chloroform/Isoamylalkohol
PCR	Polymerasekettenreaktion
PGK	Phosphoglyceratkinase
PRK	Phosphoribulokinase
PYK	Pyruvatkinase
rbcL	große Untereinheit der Ribulose-1,5-diphosphat-Carboxylase
RPE	Ribulose-5-phosphat Epimerase
rpm, Upm	Umdrehungen pro Minute
RT	Raumtemperatur
RT-PCR	Reverse Transkriptase Polymerase Ketten Reaktion
Rubisco	Ribulose-1,5-diphosphat-Carboxylase
SBP	Sedoheptulose-1,7-bisphosphatase

9. Anhang

SDS	Natriumdodecylsulfat
sec	Sekunde(n)
Tab.	Tabelle
TAE	Tris/Acetat/EDTA-Puffer
TE	Tris/EDTA-Lösung
TKL	Transketolase
TPI	Triosephosphat Isomerase
Tris/HCl	Tris-(hydroxymethyl-)aminomethan
ü.N.	über Nacht
UV	Ultraviolett
Vol	Volumen
X-Gal	5-Brom-4-chlor-indolyl- β -D-galactopyranosid

Danksagung

Mein besonderer Dank gilt Herrn Dr. Jörn Petersen, der mir stets mit Rat, Tat und konstruktiver Kritik zur Seite stand und mich beim Badminton meist „gewinnen ließ“.

Ich danke meinem Doktorvater Herrn Prof. Dr. Rüdiger Cerff für die Betreuung meiner Arbeit und die Bereitstellung des Arbeitsplatzes.

Herrn Prof. Dr. Ralf-Rainer Mendel danke ich für sein Interesse an meiner Arbeit und die bereitwillige Übernahme des Koreferats. Zudem fühlten sich meine Algen erst in seinem Kulturraum richtig wohl.

Weiterhin möchte ich mich bei Dr. Jörn Petersen, meinen Eltern und Torsten Heidenreich dafür bedanken, dass sie sich die Zeit genommen haben, die vorliegende Arbeit auf Fehler zu überprüfen.

Den Mitarbeitern des Instituts für Genetik danke ich für viele wertvolle Hinweise im Bezug auf die praktische Arbeit, anregende Diskussionen sowie das positive Arbeitsklima. Insbesondere Ulrike Brandt und Sarina Scharbatke sind hierbei aufgrund ihrer hervorragenden technischen Assistenz hervorzuheben.

An dieser Stelle möchte ich zudem meiner gesamten Familie Danke sagen, die mich während des Biologiestudiums und der Promotion in jeder Hinsicht unterstützte und durch diesen Rückhalt einen entscheidenden Beitrag zu dieser Arbeit geleistet hat.

Lebenslauf

Name: René Teich
Geburtsdatum: 10. Oktober 1977
Geburtsort: Salzwedel
Staatsangehörigkeit: deutsch
Familienstand: ledig

Schulbildung

1984 – 1991 Polytechnische Oberschule (POS) Jenny Marx in Salzwedel
1991 – 1996 Käthe-Kollwitz Gymnasium in Salzwedel
Abschluss: Allgemeine Hochschulreife

Zivildienst

09/1996 – 09/ 1997 Psychiatrisches Pflegeheim Dr. Nowak GmbH in Salzwedel

Studium

10/1997 – 08/2002 Studium der Biologie an der TU Braunschweig
09/1999 Vordiplom
08/2002 Abschluss: Diplom
Diplomarbeit am Institut für Genetik der TU Braunschweig
Thema der Diplomarbeit:
„Klonierung und Charakterisierung der Gene für cytosolische Glucose-6-phosphat Isomerase und cytosolisch/plastidäre Isoformen der Glycerinaldehyd-3-phosphat Dehydrogenase aus Haptophyten (*Pavlova lutheri* und *Prymnesium parvum*)“

Dissertation

seit 09/2002 Wissenschaftlicher Mitarbeiter am Institut für Genetik der TU Braunschweig im Rahmen einer durch die Graduiertenförderung der TU Braunschweig und der DFG finanzierten Promotionsarbeit